

普通高等学校信息与计算科学专业系列丛书

微分方程数值解法（第四版）

李荣华　刘　播

高等教育出版社
HIGHER EDUCATION PRESS



本书第二版荣获

第二届国家教委高等学校优秀教材二等奖



ISBN 978-7-04-024863-0

9 787040 248630 >

定价 25.00 元

普通高等学校信息与计算科学专业系列丛书

微分方程数值解法

（第四版）

李荣华　刘　播



高等教育出版社

内容提要

本书是编者在《微分方程数值解法》(第三版) 的基础上修订而成的。本次修订的宗旨是加强方法及其应用, 考虑到不同院校的需要, 仍然保留常微分方程数值解法这一章。为了更方便教学, 采取先介绍有限差分法, 后介绍 Galerkin 有限元法, 去掉原来的第七章, 将离散方程的有关解法与椭圆方程的差分法和有限元法合并, 同时增设了一些数值例子, 适当删减部分理论内容, 突出应用, 降低难度。本书包括六章, 第一章为常微分方程数值解法, 第二章至第四章为椭圆、抛物和双曲偏微分方程的有限差分法, 第五章、第六章为 Galerkin 有限元法。

本书是为信息与计算科学专业编写的教材, 也可以作为数学与应用数学、力学及某些工程科学专业的教学用书, 对于从事科学技术、工程与科学计算的专业人员也有参考价值。

图书在版编目(CIP)数据

微分方程数值解法/李荣华,刘播 .—4 版 .—北京:
高等教育出版社,2009.1
ISBN 978-7-04-024863-0

Ⅰ.微… Ⅱ.①李…②刘… Ⅲ.微分方程-数值计算-高等学校-教材 Ⅳ.O241.8

中国版本图书馆 CIP 数据核字(2008)第 191694 号

出版发行	高等教育出版社	购书热线	010-58581118
社　　址	北京市西城区德外大街 4 号	免费咨询	800-810-0598
邮政编码	100120	网　　址	http://www.hep.edu.cn
总　　机	010-58581000		http://www.hep.com.cn
		网上订购	http://www.landraco.com
经　　销	蓝色畅想图书发行有限公司		http://www.landraco.com.cn
印　　刷	北京印刷一厂	畅想教育	http://www.widedu.com
		版　　次	1980 年 3 月第 1 版
开　　本	787×960　1/16		2009 年 1 月第 4 版
印　　张	18.25	印　　次	2009 年 1 月第 1 次印刷
字　　数	340 000	定　　价	25.00 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

物料号　24863-00

前　言

微分方程数值解法在数值分析中占有重要的地位, 它以逼近论、数值代数等学科为基础, 反过来又推动这些学科向前发展。微分方程数值解法在科学计算、工程技术等领域有极其广泛的应用。自上世纪 40 年代以来, 它已发展成一门庞大的计算技术学科, 并早已列为原来计算数学和应用数学专业的基础课之一。与此同时, 国内外出版了不少有关专著和教材。在我国, 编者受上海理科教材会议 (1977 年) 委托, 于 1980 年与冯果忱合作为计算数学专业编写出版过第一本教材《微分方程数值解法》, 1989 年修改后出版了第二版, 1996 年经编者较大修改后又出版了第三版。1998 年高校专业目录有了调整, 原计算数学专业更名为信息与计算科学专业, 教学计划和内容也有些改变 (如有些院校建议将常微分方程数值解法合并到逼近论中去讲)。编者根据新情况, 特别是微分方程数值解法的新进展, 于 2005 年编写出版了《偏微分方程的数值解法》(高等教育出版社)。原以为这就可以满足本专业的需要了, 但实际上, 设有信息与计算科学专业的院校分布面很广, 不少院校, 主要是一些非理科院校仍采用《微分方程数值解法》(第三版), 将常微分方程和偏微分方程的数值解法并为一门课讲授。

第三版出版以来, 又过去了十二年, 笔者在教学实践中感到这个版本仍有不少缺点。例如原书在应用方面强调不够, 特别是缺少说明方法应用的例子。再如原书内容虽已做过精选, 但一些院校反映内容仍然偏多, 给教学带来一定困难。我希望趁这次再版机会, 对这些问题尽可能加以补正和调整。首先, 选材以方法为主, 指出方法如何在实际中应用, 并有针对性地选编了一些数值应用例子。为此我们还参考了 J. W. Thomas 的书: Numerical Partial Differential Equations (1995)。其次, 对原书的理论部分 (如收敛性和误差估计) 做了适当删减, 这些内容在原书中也不属于必学范围。第三, 在体系上我们也做了较大变动, 将差分法放在 Galerkin 有限元法前面, 删去原书第七章离散化方程的解法, 将主要解法与椭圆方程差分法及有限元法各章合并。这样调整后也许更便于教学。

本书选材的基本原则是少而精和可接受性, 力求选材基本, 以及涵盖对本学科发展有重要影响的内容。由于我国开设信息与计算科学专业的院校很多, 各院校的情况和要求差别又很大, 所以教师在讲授时可根据情况适当删减部分内容, 除文中打星号的节外, 其他部分也可进一步精减, 只要无损于书的整体结构。

这次修订编者虽然付出了不少努力, 但一定还有不少缺点甚至错误, 望广大师生和读者指正。最后我要感谢帮助过本书成书的同志。大连理工大学吴微同

志从工科角度提出过一些有益意见。数值例子是由吉林大学宫成春同志，研究生廖丰恒、吕超、黄冬冬等计算完成的。研究生郭玉坤帮我把手稿输入计算机。高等教育出版社张长虹编辑为本书出版和编辑加工付出了辛勤劳动。谨此对以上各位同志表示谢意!

李荣华

2008 年 8 月 22 日于长春



目　　录



第一章　常微分方程初值问题的数值解法

§1　引　　论

1.1　一阶常微分方程初值问题

设 $f(t,u)$ 在区域 $G: 0 \leqslant t \leqslant T, |u| < \infty$ 上连续, 求 $u = u(t)$ 满足

$$\frac{\mathrm{d}u}{\mathrm{d}t} = f(t,u), 0 < t \leqslant T, \tag{1.1.1$_1$}$$

$$u(0) = u_0, \tag{1.1.1$_2$}$$

其中 u_0 是给定的初值, 这就是一阶常微分方程的初值问题. 为使问题 $(1.1.1)_{1-2}$ 的解存在、唯一且连续依赖初值 u_0, 即初值问题 $(1.1.1)_{1-2}$ 适定, 还必须对右端 $f(t,u)$ 加适当限制, 通常要求 f 关于 u 满足 Lipschitz 条件: 存在常数 L, 使

$$|f(t,u_1) - f(t,u_2)| \leqslant L\,|u_1 - u_2| \tag{1.1.2}$$

对所有 $t \in [0,T]$ 和 $u_1, u_2 \in (-\infty, +\infty)$ 成立 (参看 [1]). 本章总假定 f 满足上述条件.

虽然初值问题 $(1.1.1)_{1-2}$ 对很大一类右端函数有解, 但求出所需的解绝非易事. 事实上, 除了极特殊情形外, 人们不可能求出它的精确解, 只能用各种近似方法得到满足一定精度的近似解. 读者在常微分方程教程中已经熟悉了级数解法和 Picard 逐步逼近法, 这些方法可以给出解的近似表达式, 称为近似解析方法. 另一类近似方法只给出解在一些离散点上的近似值, 称为数值方法. 由于后一类方法应用范围更广, 特别适合用计算机计算, 所以本章只讨论初值问题的数值解法.

1.2　Euler 法

最简单的数值解法是 Euler 法. 将区间 $[0,T]$ 作 N 等分, 小区间的长度 $h = T/N$ 称为步长, 点列 $t_n = nh(n = 0,1,\cdots,N)$ 称为节点, $t_0 = 0$. 由已知初值 $u(t_0) = u_0$, 可算出 $u(t)$ 在 $t = t_0$ 的导数值 $u'(t_0) = f(t_0, u(t_0)) = f(t_0, u_0)$. 利

用 Taylor 展式

$$u(t_1)=u(t_0+h)=u(t_0)+hu'(t_0)+\frac{h^2}{2}u''(t_0)+\frac{h^3}{6}u'''(\zeta)$$
$$=u_0+hf(t_0,u_0)+R_0, \tag{1.1.3}$$

其中 $\zeta\in(t_0,t_1)$, 并略去二阶小量 R_0, 得

$$u_1=u_0+hf(t_0,u_0).$$

u_1 就是 $u(t_1)$ 的近似值. 利用 u_1 又可算出 u_2, 如此下去可算出 u 在所有节点上的近似值, 一般递推公式为

$$u_{n+1}=u_n+hf(t_n,u_n),\quad n=0,1,\cdots,N-1, \tag{1.1.4}$$

这就是 Euler 法.

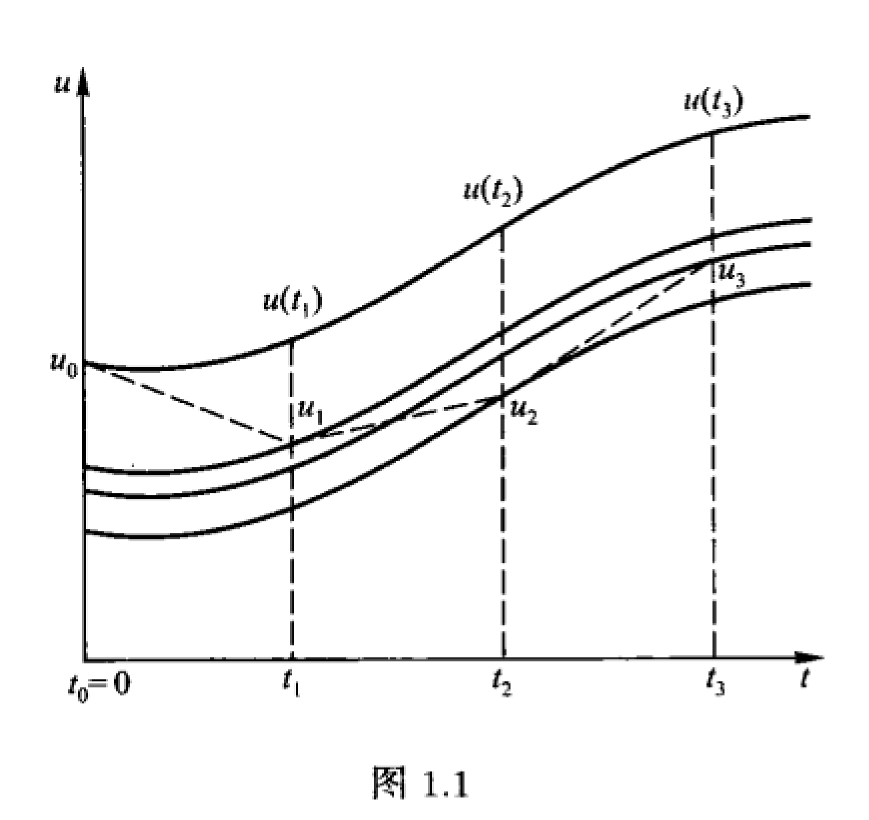


图 1.1

Euler 法有明显的几何意义. 实际上, $(1.1.1)_1$ 的解是 t,u 平面上的积分曲线族, 过任一点恰有一积分曲线通过. 按 Euler 法, 过初始点 (t_0,u_0) 作经过此点的积分曲线的切线 (斜率为 $f(t_0,u_0)$), 沿切线取点 (t_1,u_1) (u_1 按 (1.1.4) 计算) 作为 $(t_1,u(t_1))$ 的近似, 然后过 (t_1,u_1) 做一经过此点的积分曲线的切线, 沿切线取点 (t_2,u_2)(u_2 按 (1.1.4) 计算) 作为 $(t_2,u(t_2))$ 的近似. 如此下去. 即得一以 (t_n,u_n) 为顶点的折线, 这就是用 Euler 法得到的近似积分曲线 (图 1.1 中的虚折线). 从几何上看, h 越小, 此折线逼近积分曲线越好, 因此也称 Euler 法为 Euler 折线法.

现在用数值积分法推导 Euler 法. 将问题 $(1.1.1)_{1-2}$ 写成等价的积分形式:

$$u(t)=u_0+\int_{t_0}^{t}f(\tau,u(\tau))\mathrm{d}\tau\quad(t_0=0), \tag{1.1.5}$$

特别

$$u(t_1) = u_0 + \int_{t_0}^{t_1} f(\tau, u(\tau))\mathrm{d}\tau \quad (t_0 = 0).$$

用左矩形公式近似右端积分, 并用 u_1 代替 $u(t_1)$ 即得 $u_1 = u_0 + hf(t_0, u_0)$, 这就是 Euler 法 (1.1.4). 我们也可用梯形公式近似上述积分, 仍用 u_1 替代 $u(t_1)$, 得

$$u_1 = u_0 + \frac{h}{2}\left[f(t_0, u_0) + f(t_1, u_1)\right].$$

一般而言,

$$u_{n+1} = u_n + \frac{h}{2}\left[f(t_n, u_n) + f(t_{n+1}, u_{n+1})\right], n = 0, 1, \cdots, N-1, \tag{1.1.6}$$

称之为改进的 Euler 法. 显然改进的 Euler 法比 Euler 法精度更高, 但每步计算要解非线性方程 (1.1.6) (关于 u_{n+1}), 这可用如下迭代公式:

$$u_{n+1}^{[k+1]} = u_n + \frac{h}{2}\left[f(t_n, u_n) + f(t_{n+1}, u_{n+1}^{[k]})\right], \quad k = 0, 1, \cdots. \tag{1.1.7}$$

取初值为 $u_{n+1}^{[0]} = u_n$, 一般只需迭代几步即可收敛.

现在分析一下 Euler 法误差的来源. 为使问题简化, 我们不考虑因计算机字长限制引起的舍入误差. 注意 (1.1.3) 或其一般的递推式

$$u(t_{n+1}) = u(t_n) + hf(t_n, u(t_n)) + R_n \tag{1.1.8}$$

是精确方程, 其中

$$R_n = \frac{h^2}{2}u''(t_n) + \frac{h^3}{6}u'''(\zeta), \zeta \in (t_n, t_{n+1}). \tag{1.1.9}$$

由 (1.1.8) 到 Euler 法 (1.1.4) 的唯一差别是舍去了余项 R_n. 令

$$L[u_n; h] = u_{n+1} - u_n - hf(t_n, u_n), \tag{1.1.10}$$

取 $u_n = u(t_n)$, 则 $R_n = L[u(t_n); h] = u(t_{n+1}) - u(t_n) - hu'(t_n)$. 今后称 R_n 为**局部截断误差**. 显然 Euler 法的局部截断误差的阶为 $O(h^2)$.

将 $t \in [t_n, t_{n+1}]$ 表成 $t = t_n + \tau h, 0 \leqslant \tau \leqslant 1$. 由线性插值的余项公式, 我们有

$$\begin{aligned} f(t, u(t)) &= u'(t) = u'(t_n + \tau h) \\ &= u'(t_n) + \tau[u'(t_{n+1}) - u'(t_n)] + \frac{h^2}{2}\tau(\tau - 1)\, u'''(t_n + \theta h), (0 \leqslant \theta \leqslant 1) \end{aligned}$$

于是

$$\begin{aligned} \int_{t_n}^{t_{n+1}} f(t, u(t))\mathrm{d}t &= \int_0^1 \left[u'(t_n) + \tau(u'(t_{n+1}) - u'(t_n))\right] h\mathrm{d}\tau + \\ &\quad \frac{h^3}{2}\int_0^1 \tau(\tau - 1)\, u'''(t_n + \theta h)\, \mathrm{d}\tau \\ &= \frac{h}{2}\left[u'(t_n) + u'(t_{n+1})\right] - \frac{h^3}{12}u'''(\zeta), \quad \zeta \in (t_n, t_{n+1}). \end{aligned}$$

足见改进 Euler 法的局部截断误差为

$$R_n^{(1)} = -\frac{h^3}{12}u'''(\zeta), \tag{1.1.11}$$

其阶为 $O(h^3)$, 比 Euler 法高一阶.

当然我们更关心的是近似解的误差, 即

$$e_n = u(t_n) - u_n,$$

称为**整体误差**. 将 (1.1.4) 和 (1.1.8) 相减, 知 e_n 满足误差方程:

$$e_{n+1} = e_n + h\left[f(t_n, u(t_n)) - f(t_n, u_n)\right] + R_n. \tag{1.1.12}$$

因 $f(t,u)$ 关于 u 满足 Lipschitz 条件 (1.1.2), 故

$$\begin{aligned} |e_{n+1}| &\leqslant |e_n| + Lh\,|e_n| + R \\ &= (1 + Lh)\,|e_n| + R, \end{aligned}$$

其中 $R = \max\limits_n |R_n|$. 以此递推, 得

$$\begin{aligned} |e_n| &\leqslant (1+Lh)\,|e_{n-1}| + R \leqslant (1+Lh)^2|e_{n-2}| + (1+Lh)R + R \\ &\leqslant \cdots \leqslant (1+Lh)^n\,|e_0| + R\sum_{j=0}^{n-1}(1+Lh)^j \\ &= (1+Lh)^n\,|e_0| + \frac{R}{Lh}[(1+Lh)^n - 1]. \end{aligned}$$

注意 $t_n = t_0 + nh \leqslant T, n = (t_n - t_0)/h$, 于是

$$|e_n| \leqslant e^{L(T-t_0)}|e_0| + \frac{R}{Lh}(e^{L(T-t_0)} - 1), \quad n = 1, \cdots, N. \tag{1.1.13}$$

右端依赖初始误差 e_0 和局部截断误差的界 R. 对 Euler 法, 可取 $R = Ch^2$ (C 是与 n 无关的常数). 若 $e_0 = 0$ (取 $u_0 = u(t_0)$), 则

$$|e_n| \leqslant CL^{-1}e^{L(T-t_0)}h. \tag{1.1.14}$$

所以 $e_n = O(h)$, 比局部截断误差低一阶. 用同样方法可以证明改进的 Euler 法的整体误差的阶为 $O(h^2)$, 也比局部截断误差低一阶.

在实际计算中, 初值 u_0 往往不能精确给出 (例如, 包含测量误差, 舍入误差等等), 其误差将依次传递下去. 如果传递误差能够被控制, 精确说来, 传递误差连续依赖初始误差, 则说算法稳定; 否则就说不稳定. 显然不稳定的算法是不能

用的. 我们考察 Euler 法. 设从初值 u_0 和 v_0 算出的节点值分别为 $\{u_n\}$ 和 $\{v_n\}$, 则

$$u_n = u_{n-1} + hf(t_{n-1}, u_{n-1}),$$
$$v_n = v_{n-1} + hf(t_{n-1}, v_{n-1}), n = 1, 2, \cdots, N.$$

两式相减并令 $e_n = u_n - v_n$, 得

$$e_n = e_{n-1} + h\left[f(t_{n-1}, u_{n-1}) - f(t_{n-1}, v_{n-1})\right],$$

从而

$$\begin{aligned}|e_n| &\leqslant |e_{n-1}| + Lh\,|e_{n-1}| = (1+Lh)\,|e_{n-1}| \\ &\leqslant \cdots \leqslant (1+Lh)^n\,|e_0| \\ &\leqslant e^{LT}\,|e_0| \qquad (\text{因 } nh \leqslant T).\end{aligned}$$

这说明 e_n 连续依赖初始误差 e_0, 即 Euler 法稳定. 同样可证改进的 Euler 法也稳定.

1.3 线性差分方程

设 $a_0(n), a_1(n), \cdots, a_k(n)$ 和 $b_n(n = 0, 1, \cdots)$ 已知,$a_0(n) \neq 0, a_k(n) \neq 0$. 称序列 $\{u_n\}$ 满足的方程

(1.1.15) $$a_k(n)u_{n+k} + a_{k-1}(n)u_{n+k-1} + \cdots + a_0(n)u_n = b_n, n = 0, 1, 2, \cdots$$

为 k 阶线性差分方程, 序列 $\{u_n\}$ 是差分方程的解. 当右端 $b_n = 0(n = 0, 1, \cdots)$ 时, 称为齐方程. 为确定差分解, 需给定 k 个初值 $u_0, u_1, \cdots, u_{k-1}$.

记 $\Delta_+ u_n = u_{n+1} - u_n$, 称为向前差分, 则 $u_{n+1} = u_n + \Delta_+ u_n$, 即 u_{n+1} 可用 u_n 的一阶差分表示. 又二阶差分

$$\Delta_+^2 u_n = \Delta_+ u_{n+1} - \Delta_+ u_n = u_{n+2} - u_{n+1} - \Delta_+ u_n = u_{n+2} - u_n - 2\Delta_+ u_n,$$

故 $u_{n+2} = u_n + 2\Delta_+ u_n + \Delta_+^2 u_n$, 即 u_{n+2} 可用 u_n 的一阶和二阶差分表示. 依次类推, 可知 u_{n+j} 能用 u_n 的一阶、二阶直至 j 阶差分表示. 所以差分方程 (1.1.15) 的最高阶为 k. k 阶线性差分方程是 k 阶线性常微分方程的离散模拟, 二者有许多平行的基本性质. 例如:

(1) 齐方程的解具有可加性和齐次性. 若 $\{u_n\}$ 和 $\{v_n\}$ 都是齐方程的解, α 和 β 是任意常数, 则 $\{\alpha u_n + \beta v_n\}$ 也是它的解.

(2) k 阶齐方程存在 k 个线性无关的解. k 个解 $\{u_n^{(j)}\}(j = 0, 1, \cdots, k-1)$ 说是线性无关的, 如果方程

$$\sum_{j=0}^{k-1} c_j u_n^{(j)} = 0, \quad n = 0, 1, \cdots$$

仅当 $c_0 = c_1 = \cdots = c_{k-1} = 0$ 时成立. 由于任一解 $\{u_n^{(j)}\}(n = 0, 1, \cdots)$ 可表为初值 $u_0^{(j)}, \cdots, u_{k-1}^{(j)}$ 的一次组合, 所以 k 个解 $\{u_n^{(j)}\}(n = 0, 1, \cdots, k-1)$ 线性无关的充要条件是初始向量 $\left(u_0^{(j)}, \cdots, u_{k-1}^{(j)}\right)^{\mathrm{T}}$ (T 表示转置, $j = 0, 1, \cdots, k-1$) 线性无关, 即行列式

$$\begin{vmatrix} u_0^{(0)} & u_0^{(1)} & \cdots & u_0^{(k-1)} \\ u_1^{(0)} & u_1^{(1)} & \cdots & u_1^{(k-1)} \\ \vdots & \vdots & & \vdots \\ u_{k-1}^{(0)} & u_{k-1}^{(1)} & \cdots & u_{k-1}^{(k-1)} \end{vmatrix} \neq 0.$$

由此进一步推出, k 阶齐方程恰有 k 个线性无关的解, 且任一解可表示成这些解的线性组合.

(3) 非齐方程的通解等于齐方程的通解与非齐方程一特解之和.

今考虑常系数差分方程:

$$\sum_{j=0}^{k} a_j u_{n+j} = b_n, \quad n = 0, 1, \cdots, \tag{1.1.16}$$

其齐方程为

$$\sum_{j=0}^{k} a_j u_{n+j} = 0, \quad n = 0, 1, \cdots. \tag{1.1.17}$$

考虑齐方程形如 $u_n = \zeta^n$ (ζ 待定) 的解, 以之代到 (1.1.17), 知 ζ 应满足

$$a_k \zeta^k + a_{k-1}\zeta^{k-1} + \cdots + a_1\zeta + a_0 = 0,$$

即 ζ 应是代数方程

$$a_k \lambda^k + a_{k-1}\lambda^{k-1} + \cdots + a_1\lambda + a_0 = 0 \tag{1.1.18}$$

的根. 反之, 若 ζ 是 (1.1.18) 的任一根, 则 $u_n = \zeta^n$ 必为 (1.1.17) 的解. 分几种情况:

(i) 方程 (1.1.18) 有 k 个互异的实根 $\zeta_1, \zeta_2, \cdots, \zeta_k$, 则 $\zeta_1^n, \zeta_2^n, \cdots, \zeta_k^n$ 是差分方程 (1.1.17) 的 k 个线性无关解, 通解为

$$u_n = \sum_{j=1}^{k} c_j \zeta_j^n, \quad n = 0, 1, \cdots.$$

(ii) 方程 (1.1.18) 有 m 个互异实根 $\zeta_1, \zeta_2, \cdots, \zeta_m, \zeta_j$ 的重数是 $r_j, r_1 + \cdots + r_m = k$, 则

$$\zeta_j^n, n\zeta_j^n, \cdots, n^{r_j - 1}\zeta_j^n, n = 0, 1, \cdots$$

是 (1.1.17) 的 r_j 个线性无关解. (1.1.17) 的通解形如

$$u_n = \sum_{j=1}^{m}\sum_{l=1}^{r_j} c_{jl} n^{l-1} \zeta_j^n, n = 0, 1, \cdots. \tag{1.1.19}$$

(iii) 若 (1.1.18) 有复根 ζ_j, 则其共轭 $\overline{\zeta_j}$ 也是根. 令

$$\zeta_j = \rho e^{i\theta} = \rho(\cos\theta + i\sin\theta) \quad (i = \sqrt{-1}),$$

则

$$\overline{\zeta_j} = \overline{\rho e^{i\theta}} = \rho(\cos\theta - i\sin\theta).$$

此时可用两个线性无关的实解

$$\rho^n \cos n\theta, \rho^n \sin n\theta$$

替换 ζ_j^n 和 $\overline{\zeta_j^n}$.

现在给出非齐方程 (1.1.16) 的通解表达式. 引进 k 维向量 $\boldsymbol{U}_n = (u_{n+k-1}, u_{n+k-2}, \cdots, u_n)^{\mathrm{T}}$ 和 $k \times k$ 矩阵:

$$\boldsymbol{C} = \begin{bmatrix} -a_k^{-1}a_{k-1} & -a_k^{-1}a_{k-2} & \cdots & -a_k^{-1}a_1 & -a_k^{-1}a_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \tag{1.1.20}$$

将 (1.1.16) 改写成:

$$u_{n+k} = -a_k^{-1}(a_{k-1}u_{n+k-1} + \cdots + a_0 u_n) + a_k^{-1} b_n,$$

进一步写成向量形式:

$$\boldsymbol{U}_{n+1} = \boldsymbol{C}\boldsymbol{U}_n + \boldsymbol{b}_n, \quad \boldsymbol{b}_n = \left(a_k^{-1}b_n, 0, \cdots, 0\right)^{\mathrm{T}}. \tag{1.1.16$'$}$$

以此递推, 即得通解

$$\boldsymbol{U}_n = \boldsymbol{C}^n \boldsymbol{U}_0 + \sum_{l=0}^{n-1} \boldsymbol{C}^l \boldsymbol{b}_{n-l-1}, \quad n = 1, 2, \cdots, \tag{1.1.21}$$

其中第一项是齐方程的通解, 第二项是非齐方程的特解 (初值 $u_0 = u_1 = \cdots = u_{k-1} = 0$).

直接展开行列式 $|\boldsymbol{C}-\lambda\boldsymbol{I}_k|$($\boldsymbol{I}_k$ 是 k 阶单位矩阵), 即知 (1.1.18) 的左端就是 $\boldsymbol{C}$ 的特征多项式. 设 λ_j 是 $\boldsymbol{C}$ 的特征值 (方程 (1.1.18) 的根), $x_j=(d_{k-1},d_{k-2},\cdots,d_0)^{\mathrm{T}}$ 是相应的特征向量, 则

$$\begin{aligned}
-a_k^{-1}(a_{k-1}d_{k-1}+a_{k-2}d_{k-2}+\cdots+a_0d_0)&=\lambda_j d_{k-1},\\
d_{k-1}&=\lambda_j d_{k-2},\\
&\cdots\cdots\cdots\cdots\\
d_1&=\lambda_j d_0.
\end{aligned}$$

由此得

$$d_0=d_0,\quad d_1=\lambda_j d_0,\quad d_2=\lambda_j^2 d_0,\cdots,d_{k-1}=\lambda_j^{k-1}d_0,$$
$$x_j=d_0\left(\lambda_j^{k-1},\lambda_j^{k-2},\cdots,\lambda_j,1\right)^{\mathrm{T}}.$$

可见任一特征值的特征空间的维数都是 1, 因此只有单特征值的初等因子的次数才是 1. 用相似变换 $\boldsymbol{S}$ 将 $\boldsymbol{C}$ 化成 Jordan 标准型:

$$\boldsymbol{C}=\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}^{-1}.$$

与单特征值相应的 Jordan 块为 λ, 与重特征值相应的 Jordan 块为

$$\boldsymbol{J}_r=\begin{bmatrix}\lambda & 1 & & 0\\ & \ddots & \ddots & \\ & & \ddots & 1\\ 0 & & & \lambda\end{bmatrix}\quad(r:\lambda\ \text{的重数}).$$

因 $\boldsymbol{C}^n=\boldsymbol{S}\boldsymbol{J}^n\boldsymbol{S}^{-1}$, $\boldsymbol{J}^n$ 也是分块矩阵, 每一分块形如 $(\lambda)^n=(\lambda^n)$(λ 是单特征值) 或 (λ 是重特征值)

$$\boldsymbol{J}_r^n=\begin{bmatrix}\lambda^n & n\lambda^{n-1} & \cdots & n\lambda^{n-r-1}\\ & \lambda^n & \ddots & \vdots\\ 0 & & \ddots & n\lambda^{n-1}\\ & & & \lambda^n\end{bmatrix}\quad(n\geqslant r).$$

引理 1.1　(i) 矩阵族 $\{\boldsymbol{C}^n\}(n=1,2,\cdots)$ 有界的充要条件是: 方程 (1.1.18) 的所有根在单位圆内, 而位于单位圆周上的都是单根. (ii) 矩阵族 $\boldsymbol{C}^n$ 当 $n\to\infty$

时有极限的充要条件是: 方程 (1.1.18) 的所有根在单位圆内, 而位于单位圆周上的根等于 1. (iii) 矩阵族 C^n 当 $n \to \infty$ 时趋于零矩阵的充要条件是 (1.1.18) 的所有根在单位圆内部.

1.4 Gronwall 不等式

在做解的先验估计时经常要用 Gronwall 不等式 (也称 Bellman 不等式). 先介绍连续形式的 Gronwall 不等式.

引理 1.2 设连续函数 $\eta(t)(a \leqslant t \leqslant b)$ 满足

$$|\eta(t)| \leqslant \beta + \alpha \int_a^t |\eta(\tau)| \mathrm{d}\tau, \quad a \leqslant t \leqslant b, \tag{1.1.22}$$

其中 α, β 为非负常数, 则

$$|\eta(t)| \leqslant \beta \mathrm{e}^{\alpha(t-a)}, \quad a \leqslant t \leqslant b. \tag{1.1.23}$$

证明 先设 $\beta > 0$. 令

$$\zeta(t) = \beta + \alpha \int_0^t |\eta(\tau)| \mathrm{d}\tau,$$

则由 (1.1.22) 得

$$\frac{\mathrm{d}\zeta(t)}{\mathrm{d}t} \leqslant \alpha\zeta(t).$$

显然 $\zeta(t) > 0$, 故

$$\frac{\zeta'(t)}{\zeta(t)} \leqslant \alpha.$$

于 $[a,t]$ 上积分, 得

$$\ln(\zeta(t)/\beta) \leqslant \alpha\,(t-a).$$

利用 (1.1.22) 即得 (1.1.23).

今设 $\beta = 0$. 对 $\forall \delta > 0$, 由 (1.1.22) 得

$$|\eta(t)| \leqslant \delta + \alpha \int_a^t |\eta(\tau)| \mathrm{d}\tau.$$

由前面得到的不等式中令 $\delta \to 0$ 便知 $\eta(t) \equiv 0$, 故 (1.1.23) 仍成立.

现在介绍离散形式的 Gronwall 不等式.

引理 1.3 设 $\alpha, \beta \geqslant 0$ 是任意常数, 序列 $\{\eta_n\}$ 满足

$$|\eta_n| \leqslant \beta + \alpha h \sum_{j=0}^{n-1} \eta_j, \quad n = k, \quad k+1, \cdots, nh \leqslant T, \tag{1.1.24}$$

其中 $h>0$ 是步长, 则

$$|\eta_n| \leqslant e^{\alpha T}(\beta + \alpha k h M_0), \quad n \geqslant k, \quad nh \leqslant T, \tag{1.1.25}$$

其中 $M_0 = \max(|\eta_0|, |\eta_1|, \cdots, |\eta_{k-1}|)$.

证明　令 $\zeta_n = \beta + \alpha h \sum_{j=0}^{n-1} |\eta_j|$, 则 (1.1.24) 相当于

$$\zeta_n - \zeta_{n-1} \leqslant \alpha h \zeta_{n-1},$$

从而

$$\zeta_n \leqslant (1+\alpha h)\zeta_{n-1} \leqslant (1+\alpha h)^2 \zeta_{n-2} \leqslant \cdots \leqslant (1+\alpha h)^{n-k}\zeta_k \leqslant e^{\alpha T}(\beta + \alpha k h M_0),$$

于是由 (1.1.24) 得

$$|\eta_n| \leqslant \zeta_n \leqslant e^{\alpha T}(\beta + \alpha k h M_0).$$

习　　题

1. 用 Euler 法和改进的 Euler 法求 $u' = -5u (0 \leqslant t \leqslant 1), u(0) = 1$ 的数值解, 步长 $h = 0.1, 0.05$; 并比较两个算法的精度.

2. 求差分方程 $u_{n+2} - 2\mu u_{n+1} + \mu u_n = 1 (n = 0, 1, \cdots)$ 的通解, $0 < \mu < 1$. 证明 $u_n \to 1/(1-\mu)$, 当 $n \to \infty$.

3. 将 $u'' = -u (0 \leqslant t \leqslant 1)$, $u(0) = 0$, $u'(0) = 1$ 化为一阶方程组, 并用 Euler 法和改进的 Euler 法求解, 步长 $h = 0.1, 0.05$; 并比较两个算法的精度.

§2　线性多步法

用 Euler 法计算节点 $t_n = t_0 + nh (t_0 = 0)$ 的近似值 u_n 只用到前一节点的值 u_{n-1}, 所以从初值 u_0 出发可算出以后各节点的值, 这样的方法称为单步法. 为了提高解的精度, 需要构造线性多步法, 其一般形式为

$$\sum_{j=0}^{k} \alpha_j u_{n+j} = h \sum_{j=0}^{k} \beta_j f_{n+j}, \tag{1.2.1}$$

其中 $f_{n+j} = f(t_{n+j}, u_{n+j}), \alpha_j$ 和 β_j 是常数, 且 $\alpha_k \neq 0, \alpha_0$ 和 β_0 不同时为 0. 按 (1.2.1) 计算 u_{n+k} 时要用到前面 k 个节点的值 $u_n, u_{n+1}, \cdots, u_{n+k-1}$, 因此称 (1.2.1) 为多步法或 k- 步法. 又因为方程 (1.2.1) 关于 f_{n+j} 是线性的, 所以称为**线性多步法**. 为使多步法的计算能够进行, 除给定的初值 u_0 外, 还要知道附加初

值 $u_1, \cdots, u_{k-1}$, 这可用其他方法计算, 后面还要介绍. 由于多步法每计算一步用到的信息更多, 因此可望造出精度更高的算法. 若 $\beta_k = 0$, 则称方法 (1.2.1) 是显的; 若 $\beta_k \neq 0$, 则称方法 (1.2.1) 是隐式的.

2.1 数值积分法

将方程 $u' = f(t, u)$ 写成积分形式, 比如在 $[t_n, t_{n+1}]$ 上积分, 得

$$u(t_{n+1}) = u(t_n) + \int_{t_n}^{t_{n+1}} f(t, u(t))\mathrm{d}t. \tag{1.2.2}$$

适当取 $k+1$ 个节点, 以被积函数 $f(t, u(t))$ 的 k 次 Lagrange 插值多项式 $L_{n,k}(t)$ 近似代替 $f(t, u(t))$, 就可得到形如 (1.2.1) 的线性多步法. 插值节点的不同取法就导致不同的多步法.

(1) **Adams 外插法** 也称 Adams-Bashorth 法, 这是一种显式多步法. 取 $t_n, t_{n-1}, \cdots, t_{n-k}$ 为节点, 构造 f 的 Langrange 插值多项式 $L_{n,k}(t)$, 则

$$f(t, u(t)) = L_{n,k}(t) + r_{n,k}(t), \tag{1.2.3}$$

其中 $r_{n,k}(t)$ 是插值余项. 代到 (1.2.2), 得

$$u(t_{n+1}) = u(t_n) + \int_{t_n}^{t_{n+1}} L_{n,k}(t)\mathrm{d}t + \int_{t_n}^{t_{n+1}} r_{n,k}(t)\mathrm{d}t. \tag{1.2.4}$$

舍去余项

$$R_{n,k} = \int_{t_n}^{t_{n+1}} r_{n,k}(t)\mathrm{d}t, \tag{1.2.5}$$

并用 u_j 代替 $u(t_j)$, 即得

$$u_{n+1} = u_n + \int_{t_n}^{t_{n+1}} L_{n,k}(t)\mathrm{d}t. \tag{1.2.6}$$

像 Euler 法一样, 称 $R_{n,k}$ 为局部截断误差.

现在给出 (1.2.6) 的具体形式. 因为插值节点等距, 被插值点 $t \in [t_n, t_{n+1}]$ 靠近最后一个节点 t_n, 所以将 $L_{n,k}(t)$ 表示成牛顿向后插值公式更方便. 记

$$t = t_n + \tau h, \quad \tau \in [0, 1],$$

则牛顿向后插值公式为

$$\begin{aligned} L_{n,k}(t) &= L_{n,k}(t_n + \tau h) \\ &= f_n + \frac{\tau}{1!}\Delta_+ f_{n-1} + \frac{\tau(\tau+1)}{2!}\Delta_+^2 f_{n-2} + \cdots + \\ &\quad \frac{\tau(\tau+1)\cdots(\tau+k-1)}{k!}\Delta_+^k f_{n+k}, \end{aligned} \tag{1.2.7}$$

式中 Δ_+^j 表示 j 阶向前差分, $f_{n-j}=f(t_{n-j},u_{n-j})$. 引进记号

$$\binom{s}{j}=\frac{s(s-1)\cdots(s-j+1)}{j!},\quad \binom{s}{0}=1, \tag{1.2.8}$$

则

$$L_{n,k}(t)=\sum_{j=0}^{k}(-1)^j\binom{-\tau}{j}\Delta_+^j f_{n-j}. \tag{1.2.9}$$

以之代到 (1.2.6), 得

$$u_{n+1}=u_n+h\sum_{j=0}^{k}a_j\Delta_+^j f_{n-j}, \tag{1.2.10}$$

其中

$$a_j=(-1)^j\int_0^1\binom{-\tau}{j}\mathrm{d}\tau,\quad j=0,\quad 1,\cdots,k. \tag{1.2.11}$$

令 $\eta_j=hf_j$, 则还可写成

$$u_{n+1}=u_n+\sum_{j=0}^{k}a_j\Delta_+^j\eta_{n-j}. \tag{1.2.12}$$

这就是 Adams 外插公式. 显然 $k=0$ 时就是 Euler 法.

为了计算 a_j, 我们给出联系这些系数的递推公式. 定义生成函数

$$G(t)=\sum_{j=0}^{\infty}a_jt^j,$$

其中 a_j 同 (1.2.11), $j=1,2,\cdots$. 因为 a_j 的界不超过 1, 故级数当 $|t|<1$ 时收敛. 将 (1.2.11) 代到上式, 则

$$\begin{aligned}G(t)&=\sum_{j=0}^{\infty}(-1)^j\int_0^1\binom{-\tau}{j}t^j\mathrm{d}\tau=\int_0^1\sum_{j=0}^{\infty}(-t)^j\binom{-\tau}{j}\mathrm{d}\tau\\&=\int_0^1(1-t)^{-\tau}\mathrm{d}\tau=-\frac{t}{(1-t)\ln(1-t)},\end{aligned}$$

从而

$$-\frac{\ln(1-t)}{t}G(t)=\frac{1}{1-t}.$$

两端展成幂级数, 上式就是

$$\left(1+\frac{1}{2}t+\frac{1}{3}t^2+\cdots\right)(a_0+a_1t+a_2t^2+\cdots)=1+t+t^2+\cdots.$$

比较 t^n 的系数, 则得递推式:

$$a_n+\frac{1}{2}a_{n-1}+\frac{1}{3}a_{n-2}+\cdots+\frac{1}{n+1}a_0=1,\quad n=0,1,\cdots.$$

由此可依次算出系数 a_j. 表 2.1 给出前几个系数值.

表 2.1 Adams 外插公式系数值

j	0	1	2	3	4	5	6
a_j	1	$\frac{1}{2}$	$\frac{5}{12}$	$\frac{3}{8}$	$\frac{251}{720}$	$\frac{95}{288}$	$\frac{10\,987}{60\,480}$

回想插值公式的余项为

$$r_{n,k}(t)=r_{n,k}(t_n+\tau h)=(-1)^{k+1}\binom{-\tau}{k+1}h^{k+1}u^{(k+2)}(\bar{\xi}),\tag{1.2.13}$$

其中 $t_{n-k}\leqslant\bar{\xi}\leqslant t_{n+1}$, 则知

$$\begin{aligned}R_{n,k}&=h^{k+2}\int_0^1(-1)^{k+1}\binom{-\tau}{k+1}u^{(k+2)}(\bar{\xi})\mathrm{d}\tau\\&=a_{k+1}h^{k+2}u^{(k+2)}(\xi),t_{n-k}\leqslant\xi\leqslant t_{n+1}.\end{aligned}\tag{1.2.14}$$

这里用到了积分第二中值公式. 由 (1.2.14) 知道 Adams 外插法 (1.2.6) 或 (1.2.10) 的局部截断误差的阶为 $O(h^{k+2})$.

实际计算时, 常常将 (1.2.10) 右端的差分表为 f_{n-j} 的线性组合. 为此利用差分公式

$$\Delta_+^j f_{n-j}=\sum_{l=0}^{j}(-1)^l\binom{j}{l}f_{n-l},\tag{1.2.15}$$

则 (1.2.10) 化为

$$u_{n+1}=u_n+h\sum_{l=0}^{k}b_{kl}f_{n-l},\tag{1.2.16}$$

其中

$$b_{kl}=(-1)^l\sum_{j=l}^{k}a_j\binom{j}{l}.\tag{1.2.17}$$

利用表 2.1, 可给出 b_{kl} 如表 2.2.

表 2.2 系 数 b_{kl} 值

l	0	1	2	3	4	5
b_{0l}	1					
$2b_{1l}$	3	−1				
$12b_{2l}$	23	−16	5			
$24b_{3l}$	55	−59	37	−9		
$720b_{4l}$	1 901	−2 774	2 616	−1 274	251	
$1\,440b_{5l}$	4 277	−7 923	9 982	−7 298	2 877	−475

例如 $k=0,1,2,3$ 的外插公式为

$$k=0:\ u_{n+1}=u_n+hf(t_n,u_n),$$
$$k=1:\ u_{n+1}=u_n+\frac{h}{2}(3f_n-f_{n-1}),$$
$$k=2:\ u_{n+1}=u_n+\frac{h}{12}(23f_n-16f_{n-1}+5f_{n-2}),$$
$$k=3:\ u_{n+1}=u_n+\frac{h}{24}(55f_n-59f_{n-1}+37f_{n-2}-9f_{n-3}).$$

(2) **Adams 内插法** 也称 Adams-Moulton 法, 这是一种隐式多步法. 现在取插值节点为 $t_{n-k},t_{n-k+1},\cdots,t_n,t_{n+1}$ (比外插法多取一点 t_{n+1}), 构造 $u'(t)$ 或 $f(t,u(t))$ 的 $k+1$ 次 Lagrange 插值多项式 $L_{n,k}^{(1)}(t)$, 插值余项为 $r_{n,k}^{(1)}(t)$, 则

$$f=L_{n,k}^{(1)}(t)+r_{n,k}^{(1)}(t).$$

以之代到 (1.2.2) 右端, 得

$$u(t_{n+1})=u(t_n)+\int_{t_n}^{t_{n+1}}L_{n,k}^{(1)}(t)\mathrm{d}t+\int_{t_n}^{t_{n+1}}r_{n,k}^{(1)}(t)\mathrm{d}t. \tag{1.2.18}$$

舍去余项

$$R_{n,k}^{(1)}=\int_{t_n}^{t_{n+1}}r_{n,k}^{(1)}(t)\mathrm{d}t, \tag{1.2.19}$$

并用 u_j 代替 $u(t_j)$, 则得 Adams 内插法:

$$u_{n+1}=u_n+\int_{t_n}^{t_{n+1}}L_{n,k}^{(1)}(t)\mathrm{d}t. \tag{1.2.20}$$

当 $k=0$ 时, 就是改进的 Euler 法. 余项 $R_{n,k}^{(1)}$ 是内插法的局部截断误差.

现在将 (1.2.20) 具体化. 仍用牛顿向后插值公式

$$L_{n,k}^{(1)}(t)=L_{n,k}^{(t)}(t)=\sum_{j=0}^{k+1}(-1)^j\binom{-\tau}{j}\Delta_+^j f_{n-j+1}, \tag{1.2.21}$$

其中 $\tau \in [-1,0]$, 而二项系数

$$(1.2.22)\quad \binom{-\tau}{j} = \frac{-\tau(-\tau-1)\cdots(-\tau-j+1)}{j!} = (-1)^j \frac{\tau(\tau+1)\cdots(\tau+j-1)}{j!}.$$

将 (1.2.21) 代到 (1.2.20) 右端, 则得 Adams 内插公式:

$$(1.2.23)\quad u_{n+1} = u_n + h\sum_{j=0}^{k+1} a_j^* \Delta_+^j f_{n-j+1} = u_n + \sum_{j=0}^{k+1} a_j^* \Delta_+^j \eta_{n-j+1},$$

其中 $\eta_{n-j+1} = hf_{n-j+1} = hf(t_{n-j+1}, u_{n-j+1})$, 而

$$(1.2.24)\quad a_j^* = (-1)^j \int_{-1}^{0} \binom{-\tau}{j} \mathrm{d}\tau, \quad j = 0,1,\cdots,k+1.$$

用生成函数法可导出系数 a_j^* 的递推公式. 表 2.3 给出 a_j^* 的前几个值.

表 2.3 系数 a_j^* 的值

j	0	1	2	3	4	5	6
a_j^*	1	$-\frac{1}{2}$	$-\frac{1}{12}$	$-\frac{1}{24}$	$-\frac{19}{720}$	$-\frac{3}{160}$	$-\frac{863}{60\,480}$

利用插值余项公式

(1.2.25)

$$r_{n,k}^{(1)}(t) = r_{n,k}^{(1)}(t_{n+1}+\tau h) = (-1)^k \binom{-\tau}{k+2} h^{k+2} u^{(k+3)}(\overline{\xi}), \quad t_{n-k} \leqslant \overline{\xi} \leqslant t_{n+1},$$

则得

$$(1.2.26)\quad R_{n,k}^{(1)} = a_{k+2}^* h^{k+3} u^{(k+3)}(\xi), \quad t_{n-k} \leqslant \xi \leqslant t_{n+1}.$$

这里用到了积分第二中值公式. 由此可见, Adams 内插法的局部截断误差的阶为 $O(h^{k+3})$.

利用差商公式 (1.2.15), 可将 (1.2.23) 写成便于计算的形式:

$$(1.2.27)\quad u_{n+1} = u_n + h\sum_{l=0}^{k+1} b_{k+1,l}^* f_{n-l+1},$$

其中

$$(1.2.28)\quad b_{k+1,l}^* = (-1)^l \sum_{j=l}^{k+1} a_j^* \binom{j}{l}.$$

表 2.4 列出 $b_{k+1,l}^*$ 的值.

表 2.4 系数 $b^*_{k+1,l}$ 值

l	0	1	2	3	4	5
b^*_{0l}	1					
$2b^*_{1l}$	1	1				
$12b^*_{2l}$	5	8	−1			
$24b^*_{3l}$	9	19	−5	1		
$720b^*_{4l}$	251	646	−264	106	−19	
$1\,440b^*_{5l}$	475	1 427	−798	482	−173	27

例如 $k=0,1,2,3$ 的内插公式为

$$
\begin{aligned}
&k=0: u_{n+1}=u_n+hf_{n+1},\\
&k=1: u_{n+1}=u_n+\frac{h}{2}\left(f_{n+1}+f_n\right),\\
&k=2: u_{n+1}=u_n+\frac{h}{12}\left(5f_{n+1}+8f_n-f_{n-1}\right),\\
&k=3: u_{n+1}=u_n+\frac{h}{24}\left(9f_{n+1}+19f_n-5f_{n-1}+f_{n-2}\right).
\end{aligned}
$$

Adams 外插法和内插法有以下几点区别. 第一, 从表 2.1 和表 2.3 (表 2.2 和表 2.4) 知道, 按绝对值系数 a^*_j 比 a_j 小 ($b^*_{k+1,l}$ 比 b_{kl} 小), 因此计算中内插法的舍入误差影响比外插法小. 第二, 用外插法和内插法计算 t_{n+1} 处的值 u_{n+1}, 用到的已知量相同 ($k+1$ 个值 $u_n, u_{n-1}, \cdots, u_{n-k}$), 但内插法局部截断误差的阶为 $O(h^{k+3})$, 外插法局部截断误差的阶为 $O(h^{k+2})$, 前者比后者高一阶. 所以为达到相同的误差阶, 内插法比外插法可少用一个初始已知量. 第三, 外插法是显式, 计算 u_{n+1} 是直接的. 内插法是隐式, 计算 u_{n+1} 需要解方程 (1.2.27), 通常用如下迭代求解:

$$
(1.2.29)\quad u^{[m+1]}_{n+1}=u_n+hb^*_{k+1,0}f(t_{n+1},u^{[m]}_{n+1})+h\sum_{l=1}^{k+1}b^*_{k+1,l}f_{n-l+1},\quad m=0,1\cdots.
$$

当 h 充分小后, 可使

$$
\left|hb^*_{k+1,0}\frac{\partial f(t_{n+1},u_{n+1})}{\partial u_{n+1}}\right|\leqslant hb^*_{k+1,0}L<1,
$$

此时迭代 (1.2.29) 收敛. 初始近似可用外插法给出, 即

$$
(1.2.30)\qquad u^{[0]}_{n+1}=u_n+h\sum_{l=0}^{k}b_{kl}f_{n-l}.
$$

由于这是好的近似, 所以收敛是很快的, 通常迭代 2 至 3 次就可收敛.

Adams 外插法和内插法是这样得到的, 先将初值问题改写成积分形式 (1.2.2), 再用适当的数值积分离散 (1.2.2). 其实, 也可将初值问题写成其他积分形式, 例如:

$$u(t_{n+2}) - u(t_n) = \int_{t_n}^{t_{n+2}} f(t, u(t))\mathrm{d}t,$$

再用适当的数值积分代替右端积分, 例如用 Simpson 公式, 得到

$$u_{n+2} - u_n = \frac{h}{3}(f_n + 4f_{n+1} + f_{n+2}), \tag{1.2.31}$$

这是线性二步法.

还应指出, 用数值积分法只能构造一类特殊的多步法, 其系数

$$\alpha_k = 1, \quad \alpha_{k-m} = -1, \quad \alpha_l = 0,\ \text{当}\ l \neq k-m, k.$$

下面介绍更一般的待定系数法.

2.2 待定系数法

令

$$L[u(t); h] = \sum_{j=0}^{k} [\alpha_j u(t+jh) - h\beta_j u'(t+jh)]. \tag{1.2.32}$$

设 $u(t)$ 是初值问题的解, 将 $u(t+jh)$ 和 $u'(t+jh)$ 在点 t 用 Taylor 公式展开, 按 h 的同次幂合并同类项, 得

$$L[u(t); h] = c_0 u(t) + c_1 h u^{(1)}(t) + \cdots + c_q h^q u^{(q)}(t) + \cdots, \tag{1.2.33}$$

其中

$$\begin{cases} c_0 = \alpha_0 + \alpha_1 + \cdots + \alpha_k, \\ c_1 = \alpha_1 + 2\alpha_2 + \cdots + k\alpha_k - (\beta_0 + \beta_1 + \cdots + \beta_k), \\ \cdots\cdots\cdots\cdots \\ c_q = \dfrac{1}{q!}(\alpha_1 + 2^q\alpha_2 + \cdots + k^q\alpha_k) - \\ \qquad \dfrac{1}{(q-1)!} \cdot (\beta_1 + 2^{q-1}\beta_2 + \cdots + k^{q-1}\beta_k), q = 2, 3, \cdots. \end{cases} \tag{1.2.34}$$

若 $u(t)$ 有 $p+2$ 次连续微商, 则可选取 k (足够大) 和 α_j, β_j 使 $c_0 = c_1 = \cdots =$

$c_p = 0$, 而 $c_{p+1} \neq 0$, 即选 α_j, β_j 满足

$$(1.2.35)\quad \begin{cases} \alpha_0 + \alpha_1 + \cdots + \alpha_k = 0, \\ \alpha_1 + 2\alpha_2 + \cdots + k\alpha_k - (\beta_0 + \beta_1 + \cdots + \beta_k) = 0, \\ \cdots\cdots\cdots\cdots \\ \dfrac{1}{p!}(\alpha_1 + 2^p\alpha_2 + \cdots + k^p\alpha_k) - \\ \qquad \dfrac{1}{(p-1)!}\cdot(\beta_1 + 2^{p-1}\beta_2 + \cdots + k^{p-1}\beta_k) = 0. \end{cases}$$

此时

$$(1.2.36)\qquad L[u(t);h] = c_{p+1}h^{p+1}u^{(p+1)}(t) + O(h^{p+2}).$$

由于 $u'(t) = f(t, u(t))$, 则

$$(1.2.37)_1\qquad \sum_{j=0}^{k}[\alpha_j u(t_n + jh) - h\beta_j f(t_n + jh, u(t_n + jh))] = R_{n,k},$$

$$(1.2.37)_2\qquad R_{n,k} = c_{p+1}h^{p+1}u^{(p+1)}(t_n) + O(h^{p+2}).$$

略去余项 $R_{n,k}$, 并用 u_{n+j} 代替 $u(t_n + jh)$, 用 f_{n+j} 记 $f(t_{n+j}, u_{n+j})$, 就得到线性多步法 (1.2.1), 其局部截断误差 $R_{n,k} = O(h^{p+1})$. 往后将证明方法的整体误差的阶是 $O(h^p)$, 所以称此法为 **p 阶 k 步法**. 显然 p 的大小和 k 有关.

因为多步法 (1.2.1) 可以差一个非零乘数, 所以不妨设 $\alpha_k = 1$. 当 $\beta_k = 0$ 时 u_{n+k} 可用 $u_{n+k-1}, \cdots, u_n$ 直接表示, 此为显方法. 反之, 当 $\beta_k \neq 0$ 时, 求 u_{n+k} 需解一个方程 (一般用迭代法), 此为隐方法. 用待定系数法构造多步法的一个基本要求, 是选取 α_j, β_j 使局部截断误差的阶尽可能高.

作为待定系数法的一个应用, 我们讨论一般的二步法. 此时 $k = 2, \alpha_2 = 1$. 记 $\alpha_0 = \alpha$, 其余四个系数 $\alpha_1, \beta_0, \beta_1, \beta_2$ 由 $c_0 = c_1 = c_2 = c_3 = 0$ 确定, 即满足方程:

$$\begin{cases} c_0 = \alpha + \alpha_1 + 1 = 0, \\ c_1 = \alpha_1 + 2 - (\beta_0 + \beta_1 + \beta_2) = 0, \\ c_2 = \dfrac{1}{2}(\alpha_1 + 4) - (\beta_1 + 2\beta_2) = 0, \\ c_3 = \dfrac{1}{6}(\alpha_1 + 8) - \dfrac{1}{2}(\beta_1 + 4\beta_2) = 0. \end{cases}$$

解之得

$$\alpha_1 = -(1 + \alpha), \quad \beta_0 = -\frac{1}{12}(1 + 5\alpha),$$

$$\beta_1 = \frac{2}{3}(1 - \alpha), \quad \beta_2 = \frac{1}{12}(5 + \alpha).$$

所以一般二步法为

$$(1.2.38)\quad u_{n+2}-(1+\alpha)u_{n+1}+\alpha u_n=\frac{h}{12}[(5+\alpha)f_{n+2}+8(1-\alpha)f_{n+1}-(1+5\alpha)f_n].$$

由 (1.2.34) 还知道

$$c_4=\frac{1}{24}(\alpha_1+16)-\frac{1}{6}(\beta_1+8\beta_2)=-\frac{1}{24}(1+\alpha),$$
$$c_5=\frac{1}{120}(\alpha_1+32)-\frac{1}{24}(\beta_1+16\beta_2)=-\frac{1}{360}(17+13\alpha).$$

所以当 $\alpha\neq-1$ 时 $c_4\neq 0$, 方法 (2.38) 是三阶二步法. 当 $\alpha=-1$ 时, $c_4=0$, 但 $c_5\neq 0$, 方法 (1.2.38) 化为

$$(1.2.39)\qquad u_{n+2}=u_n+\frac{h}{3}(f_{n+2}+4f_{n+1}+f_n),$$

这是四阶二步法, 是具有最高阶的二步法, 称为 Milne 法. 前面我们曾用 Simpson 公式导出这一算法 (见 (1.2.31)). 此外若取 $\alpha=0$, 则 (1.2.38) 为二步 Adams 内插法; 若取 $\alpha=-5$, 则 (1.2.38) 是显方法.

2.3 预估 – 校正算法

将隐式 k 步法 (1.2.1) 写成:

$$(1.2.40)\qquad u_{n+k}+\sum_{j=0}^{k-1}\alpha_j u_{n+j}=h\beta_k f(t_{n+k},u_{n+k})+h\sum_{j=0}^{k-1}\beta_j f_{n+j},$$

其中 $f_{n+j}=f(t_{n+j},u_{n+j})$. 若已求出 $u_{n+j},j=0,1,\cdots,k-1$, 则 (1.2.40) 关于 u_{n+k} 为非线性方程, 通常用如下迭代法求解:

$$(1.2.41)\quad u_{n+k}^{[m+1]}+\sum_{j=0}^{k-1}\alpha_j u_{n+j}=h\beta_k f(t_{n+k},u_{n+k}^{[m]})+h\sum_{j=0}^{k-1}\beta_j f_{n+j}, m=0,1,\cdots,$$

其中 $u_{n+k}^{[0]}$ 为给定的迭代初值. 显然若

$$h<1/L|\beta_k|,$$

L 为 f 关于 u 的 Lipschitz 常数, 初值 $u_{n+k}^{[0]}$ 又选择适当, 则迭代 (1.2.41) 收敛.

因隐式方法 (1.2.40) 每步的计算量取决于迭代 (1.2.41) 的次数, 所以选好初值 $u_{n+k}^{[0]}$ 非常重要. 最自然的一种方法是用显式多步法计算 $u_{n+k}^{[0]}$, 比如

$$(1.2.42)\qquad u_{n+k}^{[0]}+\sum_{j=0}^{k-1}\alpha_j^* u_{n+j}=h\sum_{j=0}^{k-1}\beta_j^* f_{n+j}.$$

称 (1.2.42) 为预估算式 (P 算式), (1.2.41) 为校正算式 (C 算式), 统称 (1.2.41)–(1.2.42) 为预估 – 校正算法, 简称预 – 校算法或 PC 算法 (Predictor–Corrector methods).

一个极端情形是允许 (1.2.41) 不断进行, 直至不等式 $\left|u_{n+k}^{[m+1]}-u_{n+k}^{[m]}\right|<\varepsilon$ 成立, 其中 ε 是预先指定的容许误差. 由于这种算法对迭代次数不加限制, 花费在计算函数 f 的工作量可能很大, 所以通常采用另一种限制迭代次数的算法. 假定校正次数 M (即迭代次数) 固定, P 表示预估算子, C 是一次校正算子 (即迭代一次), E 是计算 f 一次的运算, 则预估一次校正 M 次的算法可记为 $P(EC)^M=P(EC)\cdots(EC)$, 计算过程如下:

$$
\begin{aligned}
&P: u_{n+k}^{[0]}+\sum_{j=0}^{k-1}\alpha_j^* u_{n+j}^{[M]}=h\sum_{j=0}^{k-1}\beta_j^* f_{n+j}^{[M-1]},\\
&E: f_{n+k}^{[m]}=f(t_{n+k},u_{n+k}^{[m]}),\\
&C: u_{n+k}^{[m+1]}+\sum_{j=0}^{k-1}\alpha_j u_{n+j}^{[M]}=h\beta_k f_{n+k}^{[m]}+h\sum_{j=0}^{k-1}\beta_j f_{n+j}^{[M-1]},
\end{aligned}
\tag{1.2.43}
$$

其中 $m=0,1,\cdots,M-1$. 按这一预–校格式计算结束时, 得到的数据是 $u_{n+k}^{[M]}$ 和 $f_{n+k}^{[M-1]}=f(t_{n+k},u_{n+k}^{[M-1]})$, 为下一步 ($t=t_{n+1+k}$) 预估计算所用. 显然 $u_{n+k}^{[M]}$ 比 $u_{n+k}^{[M-1]}$ 更接近 u_{n+k}, 因此还可以设计一种算法, 每一步算出 $u_{n+k}^{[M]}$ 后, 利用它将 $f_{n+k}^{[M]}=f(t_{n+k},u_{n+k}^{[M]})$ 算出, 供下一步预估时使用. 这种预–校算法记为 $P(EC)^ME$, 计算过程如下:

$$
\begin{aligned}
&P: u_{n+k}^{[0]}+\sum_{j=0}^{k-1}\alpha_j^* u_{n+j}^{[M]}=h\sum_{j=0}^{k-1}\beta_j^* f_{n+j}^{[M]},\\
&E: f_{n+k}^{[m]}=f(t_{n+k},u_{n+k}^{[m]}),\\
&C: u_{n+k}^{[m+1]}+\sum_{j=0}^{k-1}\alpha_j u_{n+j}^{[M]}=h\beta_k f_{n+k}^{[m]}+h\sum_{j=0}^{k-1}\beta_j f_{n+j}^{[M]},\quad 0\leqslant m\leqslant M-1,\\
&E: f_{n+k}^{[M]}=f(t_{n+k},u_{n+k}^{[M]}).
\end{aligned}
\tag{1.2.44}
$$

原则上任一显式多步法和隐式多步法都可搭配成预 – 校算法及各种计算方案.

例 2.1 Adams 4 步 4 阶预 – 校算法. 取 4 步 4 阶 Adams 外插法为预估算法, 3 步 4 阶 Adams 内插法为校正算法, 即得

$$
\begin{aligned}
&P: u_{n+4}-u_{n+3}=\frac{h}{24}(55f_{n+3}-59f_{n+2}+37f_{n+1}-9f_n),\\
&C: u_{n+4}-u_{n+3}=\frac{h}{24}(9f_{n+4}+19f_{n+3}-5f_{n+2}+f_{n+1}).
\end{aligned}
$$

例 2.2 Milne 方法 ([24]). 以 4 步 4 阶多步法

$$P: u_{n+4} - u_n = \frac{4h}{3}\left(2f_{n+3} - f_{n+2} + 2f_{n+1}\right)$$

为预估算法, 2 步 4 阶算法

$$C: u_{n+4} - u_{n+2} = \frac{h}{3}(f_{n+4} + 4f_{n+3} + f_{n+2})$$

为校正算法, 得到由 P 和 C 组成的预–校方案 $PECE$, 称为 Milne 算法, 计算公式为

$$\begin{aligned}
&P: u_{n+4}^{[0]} - u_n^{[1]} = \frac{4h}{3}(2f_{n+3}^{[1]} - f_{n+2}^{[1]} + 2f_{n+1}^{[1]}),\\
&E: f_{n+4}^{[0]} = f(t_{n+4}, u_{n+4}^{[0]}),\\
&C: u_{n+4}^{[1]} - u_{n+2}^{[1]} = \frac{h}{3}(f_{n+4}^{[0]} + 4f_{n+3}^{[1]} + f_{n+2}^{[1]}),\\
&E: f_{n+4}^{[1]} = f(t_{n+4}, u_{n+4}^{[1]}).
\end{aligned}$$

2.4 多步法的计算问题

用 k 步法计算时, 需要知道 k 个初值 $u_0, u_1, \cdots, u_{k-1}$, 其中 $u_0 = u(t_0)$ 是给定的初值, 其余是附加初值. 计算附加初值主要是用单步法, 比如 Euler 法和 §4 将要介绍的 Runge–Kutta 法及其他单步法. 为了保持多步法的精度, 计算附加初值时要将 t_0, t_{k-1} 之间的节点加密或采用和多步法有同样阶的 Runge–Kutta 法.

多步法的第二个问题是如何选择阶 p (或者步数 k). 从收敛阶的观点, 自然希望把 p 取大一些. 但是高阶收敛方法要求解的光滑性也高, 否则达不到高精度的目的. 从后面关于绝对稳定性的分析还知道, 高阶多步法的绝对稳定域也小, 所以 p 的大小要考虑到解的光滑性和稳定性以及总的工作量.

多步法的第三个问题是步长 h 的选取. 理论上似乎按照下节的误差估计式选定 h 是合理的, 但那种估计往往偏大, 因此选定的 h 可能过小, 既不必要也不经济. 实际用的步长 h 不是一次取定, 而是根据精度要求, 由粗到细逐渐调整 (选步长), 当 h 达到要求后就以此为步长计算. 在计算中还可以改变步长, 但计算过程变复杂了, 这是多步法的缺点. 与此相反, 单步法 (如 Euler 法及后面要介绍的 Runge–Kutta 法) 则适合用变步长计算.

习 题

1. 用待定系数法求三步四阶方法类, 确定三步四阶显式方法.

2. 满足条件 $\beta_j = 0, \quad j = 0, 1, 2, \cdots, k-1$ 的 k 阶 k 步法叫做 Gear 法, 试对 $k = 1, 2, 3, 4$ 求 Gear 法的表达式.

3. 用三阶 Adams 内插法及外插法分别解初值问题 $u' = -5u$, $u(0) = 1$. 取步长 $h = 0.1, 0.05$. 观察解在 $t = 1$ 处的误差, 并与用 Euler 法计算的结果比较 (参看 §1 习题 1).

§3 相容性、稳定性和误差估计

本节讨论线性多步法的几个基本理论问题: 相容性、稳定性和误差估计.

3.1 局部截断误差和相容性

考虑初值问题

$$u' = f(t, u), \quad t \in [t_0, T] = [0, T], \tag{1.3.1$_1$}$$

$$u(t_0) = u_0 \tag{1.3.1$_2$}$$

和逼近它的 p 阶 k 步法:

$$\sum_{j=0}^{k} \alpha_j u_{n+j} = h \sum_{j=0}^{k} \beta_j f_{n+j}, \quad n = 0, 1 \cdots . \tag{1.3.2}$$

要想 (1.3.2) 的解 u_n 逼近精确解 $u(t_n)$, 必需 (1.3.2) 在某种意义下逼近 $(1.3.1)_1$. 引进差分算子

$$L[u(t); h] = \sum_{j=0}^{k} [\alpha_j u(t + jh) - h\beta_j u'(t + jh)]. \tag{1.3.3}$$

设 $u(t)$ 是 $(1.3.1)_1$ 的具有 $p + 2$ 阶连续微商的解 ($u(t) \in C^{p+2}$), 则由 (1.2.36) 和 (1.2.37), 我们有

$$\begin{aligned} L[u(t_n); h] &= \sum_{j=0}^{k} \alpha_j u(t_n + jh) - h \sum_{j=0}^{k} \beta_j u'(t_n + jh) \\ &= c_{p+1} h^{p+1} u^{(p+1)}(t_n) + O(h^{p+2}) \end{aligned} \tag{1.3.4}$$

或

$$\sum_{j=0}^{k} \alpha_j u(t_{n+j}) = h \sum_{j=0}^{k} \beta_j f(t_{n+j}, u(t_{n+j})) + L[u(t_n); h], \tag{1.3.5}$$

$$L[u(t_n); h] = c_{p+1} h^{p+1} u^{(p+1)}(t_n) + O(h^{p+2}), \tag{1.3.6}$$

其中 (参看 (1.2.34))

$$c_{p+1} = \frac{1}{(p+1)!}(\alpha_1 + 2^{p+1}\alpha_2 + \cdots + k^{p+1}\alpha_k) - \frac{1}{p!}(\beta_1 + 2^p\beta_2 + \cdots + k^p\beta_k). \tag{1.3.7}$$

像 §2 那样称 $L[u(t_n);h]$ 为**局部截断误差**, 而称 $c_{p+1}h^{p+1}u^{(p+1)}(t_n)$ 为**局部截断误差的主项**, c_{p+1} 为**误差主项系数**. 在 (1.3.5) 中舍去 $L[u(t_n);h]$, 并用 u_{n+j} 代 $u(t_{n+j})$ 就导致多步法 (1.3.2). 我们关心的是误差 $e_n = u(t_n) - u_n$, 称为**整体误差**.

现在考虑一般的 k 步法 (1.3.2) (不必要求是 p 阶方法). 为使 (1.3.2) 的解 u_n 当 $h \to 0$ 时有可能收敛到 $(1.3.1)_1$ 的解 $u(t)$, 自然要求

$$\frac{1}{h}\left[\sum_{j=0}^{k}\alpha_j u(t_{n+j}) - h\sum_{j=0}^{k}\beta_j f(t_{n+j}, u(t_{n+j}))\right] - [u'(t_n) - f(t_n, u(t_n))] = o(1)(h \to 0), \tag{1.3.8}$$

而 $u'(t) = f(t, u(t))$, 在 (1.3.3) 中令 $t = t_n$, 则 (1.3.8) 可写成:

$$L[u(t_n);h] = o(h)(h \to 0). \tag{1.3.9}$$

称多步法 (1.3.2) 相容, 如果对 $(1.3.1)_{1-2}$ 的任意光滑解 $u(t)$, 关系 (1.3.8) 或 (1.3.9) 成立. 注意当 $u(t)$ 和 $f(t,u)$ 连续可微时, (1.3.8) 右端的 $o(1) = O(h)$, 从而 $L[u(t);h] = O(h^2)$, 所以多步法 (1.3.2) 至少是一阶的 (参看 (1.2.37)). 这样可将相容性定义为

定义 3.1 解初值问题 $(1.3.1)_{1-2}$ 的多步法 (1.3.2) 说是**相容的** (Consistent), 如果它至少是一阶的.

引进多步法 (1.3.2) 的第一和第二特征多项式:

$$\rho(\lambda) = \sum_{j=0}^{k}\alpha_j\lambda^j, \tag{1.3.10}$$

$$\sigma(\lambda) = \sum_{j=0}^{k}\beta_j\lambda^j. \tag{1.3.11}$$

由 (1.2.34) 推出

定理 3.1 为使 k 步法 (1.3.2) 相容, 必需且只需

$$\rho(1) = 0, \quad \rho'(1) = \sigma(1). \tag{1.3.12}$$

3.2 稳定性

用多步法计算时, 各种因素 (如初值 $u_0, \cdots, u_{k-1}$) 是有误差的, 且这些误差将在计算中传递下去. 如果误差积累无限增长, 将会歪曲精确解, 这样的算法是不能用的, 为此我们对多步法提出稳定性要求.

定义 3.2 称多步法 (1.3.2) **稳定** (Stable), 若存在常数 C (不依赖 h 和 (1.3.2) 的解) 和 $h_0>0$, 使 $\forall h\in(0,h_0)$ 和 (1.3.2) 的任何两个解 $\{u_n\}$ 和 $\{v_n\}$(初值不同), 恒有

$$\max_{nh\leqslant T}|u_n-v_n|\leqslant C\max_{0\leqslant j<k}|u_j-v_j|. \tag{1.3.13}$$

这等于说, 对一切充分小的 h, 多步法的解连续依赖初值.

定理 3.2 设 $\rho(\lambda)$ 是形如 (1.3.10) 的第一特征多项式, 则线性多步法 (1.3.2) 稳定的充要条件是 $\rho(\lambda)$ 满足根条件, 即 $\rho(\lambda)$ 的所有根在单位圆内 ($|\lambda|\leqslant 1$), 且位于单位圆周上的根都是单根.

证明 必要性 将多步法用于方程 $u'=0(f=0)$. 此时 (1.3.2) 简化为

$$\sum_{j=0}^{k}\alpha_j u_{n+j}=0,\quad n=0,1,\cdots,$$

其通解形如 (1.1.19). 又 $\{v_n=0\}$ 是上述方程的平凡解, 不等式 (1.3.13) 化为

$$\max_{nh\leqslant T}|u_n|\leqslant C\max_{0\leqslant j<k}|u_j|, 0<h<h_0,$$

即 $\{u_n\}$ 关于 n 和 $h(nh\leqslant T, 0<h<h_0)$ 一致有界. 而当 $h\to 0$ 时 n 可趋于 ∞, 由引理 1.1 的 (i), $\rho(\lambda)$ 必满足根条件.

充分性 设 $\{u_n\}$ 和 $\{v_n\}$ 是 (1.3.2) 的任何两个解, 则

$$\sum_{j=0}^{k}\alpha_j u_{n+j}=h\sum_{j=0}^{k}\beta_j f(t_{n+j},u_{n+j}),$$

$$\sum_{j=0}^{k}\alpha_j v_{n+j}=h\sum_{j=0}^{k}\beta_j f(t_{n+j},v_{n+j}).$$

令 $e_n=u_n-v_n$, 则 e_n 满足

$$\sum_{j=0}^{k}\alpha_j e_{n+j}=hb_n,\quad n=0,1,\cdots,T/h,\quad 0<h<h_0, \tag{1.3.14}$$

其中

$$b_n=\sum_{j=0}^{k}\beta_j[f(t_{n+j},u_{n+j})-f(t_{n+j},v_{n+j})]. \tag{1.3.15}$$

设 $B=\max(|\beta_0|,|\beta_1|,\cdots,|\beta_k|)$, f 关于 u 满足 Lipschitz 条件:

$$|f(t,u)-f(t,v)|\leqslant L|u-v|,$$

则

$$|b_n| \leqslant BL\sum_{j=0}^{k}|e_{n+j}|. \tag{1.3.16}$$

如节 1.3, 引进向量 $\boldsymbol{E}_n=(e_{n+k-1},e_{n+k-2},\cdots,e_n)^{\mathrm{T}},\boldsymbol{B}_n=(h\alpha_k^{-1}b_n,0,\cdots,0)^{\mathrm{T}}$ (k 维) 和矩阵

$$\boldsymbol{C}=\begin{bmatrix}-\alpha_k^{-1}\alpha_{k-1} & -\alpha_k^{-1}\alpha_{k-2} & \cdots & -\alpha_k^{-1}\alpha_1 & -\alpha_k^{-1}\alpha_0\\ 1 & 0 & \cdots & 0 & 0\\ 0 & 1 & \cdots & 0 & 0\\ \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & 0\end{bmatrix}$$

便可将 (1.3.14) 写成向量形式:

$$\boldsymbol{E}_{n+1}=\boldsymbol{C}\boldsymbol{E}_n+\boldsymbol{B}_n,$$

进而有

$$\boldsymbol{E}_n=\boldsymbol{C}^n\boldsymbol{E}_0+\sum_{l=0}^{n-1}\boldsymbol{C}^l\boldsymbol{B}_{n-l-1},\quad n=1,2,\cdots,T/h,\quad 0<h<h_0. \tag{1.3.17}$$

今设 $\rho(\lambda)$ 满足根条件, 则由引理 1.1 的 (i), 矩阵 $\{\boldsymbol{C}^n\}$ 一致有界. 以 $||\boldsymbol{E}_n||$ 表示向量的欧氏模, $||\boldsymbol{C}||$ 表示相应的矩阵模, 则有常数 M 使

$$||\boldsymbol{C}^n|| \leqslant M,\quad n=1,2,\cdots,T/h,\quad 0<h<h_0. \tag{1.3.18}$$

又

$$||\boldsymbol{B}_n|| \leqslant h|\alpha_k^{-1}||b_n| \leqslant BL|\alpha_k^{-1}|h\sum_{j=0}^{k}|e_{n+j}|, \tag{1.3.19}$$

于是由 (1.3.17) 和 (1.3.18) 得

$$||\boldsymbol{E}_n|| \leqslant M||\boldsymbol{E}_0||+MBL|\alpha_k^{-1}|h\sum_{l=0}^{n-1}\sum_{j=0}^{k}|e_{n+j-l-1}|, \tag{1.3.20}$$

而

$$\begin{aligned}\sum_{l=0}^{n-1}\sum_{j=0}^{k}|e_{n+j-l-1}| &= \sum_{l=0}^{n-1}|e_{n+k-l-1}| + \sum_{l=0}^{n-1}\sum_{j=0}^{k-1}|e_{n+j-l-1}| \\ &\leqslant \sum_{l=0}^{n-1}||\boldsymbol{E}_{n-l}|| + \sqrt{k}\sum_{l=0}^{n-1}||\boldsymbol{E}_{n-l-1}|| \\ &\leqslant ||\boldsymbol{E}_n|| + \left(\sqrt{k}+1\right)\sum_{l=0}^{n-1}||\boldsymbol{E}_{n-l-1}|| \\ &\leqslant ||\boldsymbol{E}_n|| + \left(\sqrt{k}+1\right)\sum_{j=0}^{n-1}||\boldsymbol{E}_j||.\end{aligned}$$

故

$$||\boldsymbol{E}_n|| \leqslant M||\boldsymbol{E}_0|| + MBL|\alpha_k^{-1}|h||\boldsymbol{E}_n|| + MBL|\alpha_k^{-1}|h\left(\sqrt{k}+1\right)\sum_{j=0}^{n-1}||\boldsymbol{E}_j||$$

取 $h>0$ 充分小, 例如 $h<h_0$, 使

$$MBL|\alpha_k^{-1}|h<1.$$

令

$$\begin{aligned}K_1 &= (1-MBL|\alpha_k^{-1}|h)^{-1}M, \\ K_2 &= (1-MBL|\alpha_k^{-1}|h)^{-1}MBL|\alpha_k^{-1}|\left(\sqrt{k}+1\right),\end{aligned}$$

则

$$||\boldsymbol{E}_n|| \leqslant K_1||\boldsymbol{E}_0|| + K_2h\sum_{j=0}^{n-1}||\boldsymbol{E}_j||. \tag{1.3.21}$$

最后, 利用 Gronwall 不等式 (1.1.25) 得到

$$||\boldsymbol{E}_n|| \leqslant e^{K_2T}(K_1+K_2kh)||\boldsymbol{E}_0||, \quad n=1,2,\cdots,T/h, \quad 0<h<h_0. \tag{1.3.22}$$

这证明了多步法稳定.

对单步法 $(k=1)$: $\alpha_1u_{n+1}+\alpha_0u_n=h(\beta_1f_{n+1}+\beta_0f_n), \rho(\lambda)=\alpha_1\lambda+\alpha_0$, 唯一根 $\lambda_1=-\alpha_0/\alpha_1$. 如果方法相容, 则 $\alpha_1+\alpha_0=0, \lambda_1=1$, 且是单根, 故稳定. 特别 Euler 法稳定.

对 Admas 外插法和内插法, 相应的 $\rho(\lambda)=\lambda^k-\lambda^{k-1}=(\lambda-1)\lambda^{k-1}(k\geqslant 2)$, 除单根 $\lambda=1$ 在单位圆周上外, 其余的重根 $\lambda=0$ 都在单位圆内部, 所以 Admas

法稳定. 若 $\rho(\lambda) = \lambda^k - \lambda^{k-2} = (\lambda^2 - 1)\lambda^{k-2}(k \geqslant 2)$, 则称相应的显方法为 Nyström 法, 相应的隐式方法为广义 Milne 法. 因为唯一可能的重根 $\lambda = 0$ 在单位圆内部, 而在单位圆周上的根 $\lambda = \pm 1$ 都是单根, 所以稳定. 特别 Milne 方法 (1.2.39) 稳定. 不稳定的方法是不能用的.

例 3.1 初值问题

$$
\begin{aligned}
&u' = 4tu^{\frac{1}{2}}, \quad 0 \leqslant t \leqslant 2, \\
&u(0) = 1,
\end{aligned}
$$

的精确解为 $u(t) = (1 + t^2)^2$. 考虑线性二步法:

$$
u_{n+2} - (1 + a)u_{n+1} + au_n = \frac{1}{2}h[(3 - a)f_{n+1} - (1 + a)f_n],
$$

当 $a \neq -5$ 时是二阶方法, $a = -5$ 时是三阶方法. 第一特征多项式

$$
\rho(\lambda) = \lambda^2 - (1 + a)\lambda + a = (\lambda - 1)(\lambda - a).
$$

当 $a = 0$ 时稳定, $a = -5$ 时不稳定. 取步长 $h = 0.1$, 初值 $u_0 = 1$, 附加初值 $u_1 = (1 + h^2)^2 (h = 0.1)$ 是精确的. 用方案 (i) $a = 0$ 和 (ii) $a = -5$ 计算. 在开始不多几步, 两种方案算出的结果都和精确解符合, 且 (ii) 比 (i) 更精确. 但再往后算, 方案 (i) 的结果仍和精确解基本符合, 方案 (ii) 的误差则急剧增长, 完全歪曲了精确解, 具体数据如表 3.1(参看 [24], p.34).

表 3.1

t	精确解	(i) $a = 0$	(ii) $a = -5$
0	1.000 000 0	1.000 000 0	1.000 000 0
0.1	1.020 100 0	1.020 100 0	1.020 100 0
0.2	1.081 600 0	1.080 700 0	1.081 200 0
0.3	1.188 100 0	1.185 248 1	1.189 238 5
0.4	1.345 600 0	1.339 629 8	1.338 866 0
0.5	1.562 500 0	1.552 090 0	1.592 993 5
⋮	⋮	⋮	⋮
1.0	4.000 000 0	3.940 690 3	−68.639 804

续表

t	精确解	(i) $a=0$	(ii) $a=-5$
1.1	4.884 100 0	4.808 219 7	+367.263 92
⋮	⋮	⋮	⋮
2.0	25.000 000 0	24.632 457	-6.96×10^{8}

3.3 收敛性和误差估计

有了前述准备, 我们就可证明数值解的收敛性并估计整体误差. 设 $u(t)$ 为初值问题 $(1.3.1)_1$、$(1.3.1)_2$ 的解, u_n 是线性多步法 (1.3.2) 的解, 则 $u(t_n)$ 和 u_n 分别满足

$$\sum_{j=0}^{k}\alpha_j u(t_{n+j})=h\sum_{j=0}^{k}\beta_j f(t_{n+j},u(t_{n+j}))+L[u(t_n);h] \tag{1.3.23}$$

和

$$\sum_{j=0}^{k}\alpha_j u_{n+j}=h\sum_{j=0}^{k}\beta_j f_{n+j}\quad (f_{n+j}=f(t_{n+j},u_{n+j})). \tag{1.3.24}$$

两式相减, 则误差 (整体误差) $e_n=u(t_n)-u_n$ 满足

$$\sum_{j=0}^{k}\alpha_j e_{n+j}=hb_n+L[u(t_n);h], \tag{1.3.25}$$

$$b_n=\sum_{j=0}^{k}\beta_j[f(t_{n+j},u(t_{n+j}))-f(t_{n+j},u_{n+j})]. \tag{1.3.26}$$

如前所设, f 关于 u 满足 Lipschitz 条件, 因此 b_n 仍满足不等式 (1.3.16). 若 k 步法 (1.3.2) 相容, 例如, 设 (1.3.2) 是 p 阶方法, 则 $L[u(t_n);h]$ 有渐近表示 (1.3.6), 于是可将 (1.3.25) 写成:

$$\sum_{j=0}^{k}\alpha_j e_{n+j}=hb_n^*, \tag{1.3.27}$$

其中

$$b_n^*=b_n+c_{p+1}h^p u^{(p+1)}(t_n)+O(h^{p+1}).$$

令 $\boldsymbol{B}_n^* = (h\alpha_k^{-1}b_n^*, 0, \cdots, 0)^{\mathrm{T}}$, 则由 (1.3.16), 有

$$\begin{aligned}||\boldsymbol{B}_n^*|| &\leqslant |h\alpha_k^{-1}b_n^*| \leqslant h|\alpha_k^{-1}||b_n| + M_{p+1}h^{p+1} \\ &\leqslant hBL|\alpha_k^{-1}|\sum_{j=0}^{k}|e_{n+j}| + M_{p+1}h^{p+1} \\ &\quad (M_{p+1} = |\alpha_k^{-1}|C_{p+1}\sup_t|u^{(p+1)}(t)| + 1).\end{aligned}$$

比较 (1.3.27) 和 (1.3.14), 知向量 $\boldsymbol{E}_n = (e_{n+k-1}, \cdots, e_n)^{\mathrm{T}}$ 仍可表成 (1.3.17), 只需用 $\boldsymbol{B}_n^*$ 代替那里的 $\boldsymbol{B}_n$.

若 k 步法 (1.3.2) 稳定, 即 $\rho(\lambda)$ 满足根条件, 则 (1.3.18) 成立, 于是和 (1.3.20) 平行地有

$$||\boldsymbol{E}_n|| \leqslant M(||\boldsymbol{E}_0|| + M_{p+1}Th^p) + MBL|\alpha_k^{-1}|h\sum_{l=0}^{n-1}\sum_{j=0}^{k}|e_{n+j-l-1}|.$$

取 $h > 0$ 充分小, 使

$$MBL|\alpha_k^{-1}|h < 1.$$

则可得到与 (1.3.21) 平行的不等式:

$$||\boldsymbol{E}_n|| \leqslant K_1||\boldsymbol{E}_0 + M_{p+1}Th^p|| + K_2h\sum_{j=0}^{n-1}||\boldsymbol{E}_j||.$$

最后由 Gronwall 不等式就得到误差估计:

$$||\boldsymbol{E}_n|| \leqslant e^{K_2T}((K_1 + K_2kh)||\boldsymbol{E}_0|| + M_{p+1}Th^p). \tag{1.3.28}$$

总之我们得

定理 3.3 若解初值问题 $(1.3.1)_{1-2}$ 的多步法 (1.3.2) 相容而且稳定, 则当 $h \to 0, t_n \to t$ 时数值解 $u_n \to u(t)$, 其中 $\boldsymbol{E}_0$ 是初始误差向量. 若更设 (1.3.2) 是 p 阶方法, 则还有误差估计 (1.3.28).

习　　题

1. 证明线性多步法

$$u_{n+2} + (b-1)u_{n+1} - bu_n = \frac{1}{4}h[(b+3)f_{n+2} + (3b+1)f_n]$$

当 $b \neq -1$ 时阶是 2; 当 $b = -1$ 时阶是 3. 又 $b = -1$ 是不稳定的. 将 $b = -1$ 的方法用到 $u' = u, u(0) = 1$, 解出相应的差分方程 $(u_0 = 1, u_1 = 1)$, 说明方法发散.

2. 确定 α 的变化域, 使线性多步法

$$u_{n+3} + \alpha(u_{n+2} - u_{n+1}) - u_n = \frac{1}{2}(3+\alpha)h(f_{n+2} + f_{n+1})$$

是稳定的, 并说明方法的阶不能大于 2.

§4 单步法和 Runge – Kutta (龙格 – 库塔) 法

Euler 法是最简单的单步法. 单步法不需要附加初值, 所需的存储量小, 改变步长灵活, 但线性单步法的阶最多是 2. 本节将介绍非线性 (关于 f) 高阶单步法, 重点是 Runge – Kutta 法.

4.1 Taylor 展开法

设初值问题

$$\begin{cases} u' = f(t,u), \\ u(t_0) = u_0 \end{cases}$$

的解充分光滑. 将 $u(t)$ 在 t_0 处用 Taylor 公式展开:

$$u(t_1) = u(t_0) + hu'(t_0) + \frac{h^2}{2!}u^{(2)}(t_0) + \cdots + \frac{h^p}{p!}u^{(p)}(t_0) + O(h^{p+1}), \tag{1.4.1}$$

其中 $u(t_0) = u_0$, $u'(t_0) = f(t_0, u(t_0)) = f(t_0, u_0)$,

$$\begin{cases} u^{(2)}(t_0) = \left.\dfrac{\mathrm{d}}{\mathrm{d}t}f\right|_{t=t_0} = [f_t + f_u \cdot u'(t)]_{t=t_0} \\ \qquad\qquad = f_t(t_0,u_0) + f(t_0,u_0)f_u(t_0,u_0), \\ u^{(3)}(t_0) = \dfrac{\mathrm{d}}{\mathrm{d}t}\left[\dfrac{\mathrm{d}}{\mathrm{d}t}f\right]_{t=t_0} \\ \qquad\qquad = f_{tt}(t_0,u_0) + 2f(t_0,u_0)f_{tu}(t_0,u_0) + \\ \qquad\qquad\quad f^2(t_0,u_0)f_{uu}(t_0,u_0) + f_t(t_0,u_0)f_u(t_0,u_0) + \\ \qquad\qquad\quad f(t_0,u_0)f_u^2(t_0,u_0), \\ \qquad\qquad\quad \cdots\cdots\cdots\cdots \end{cases} \tag{1.4.2}$$

令

$$\varphi(t,u(t),h) = \sum_{j=1}^{p} \frac{h^{j-1}}{j!}\frac{\mathrm{d}^{j-1}}{\mathrm{d}t^{j-1}}f(t,u(t)), \tag{1.4.3}$$

则可将 (1.4.1) 改写为

$$u(t_0+h) - u(t_0) = h\varphi(t_0, h(t_0), h) + O(h^{p+1}). \tag{1.4.4}$$

舍去余项 $O(h^{p+1})$, 则得

$$u_1 - u_0 = h\varphi(t_0, u_0, h).$$

一般说来, 若已知 u_n, 则

$$u_{n+1} - u_n = h\varphi(t_n, u_n, h), \quad n = 0, 1, \cdots. \tag{1.4.5}$$

这是一个单步法, 局部截断误差为 $O(h^{p+1})$. 由 (1.4.2)(1.4.3) 可知 φ 关于 f 非线性. 当 $p = 1$ 时它是 Euler 折线法. 由于计算 $\varphi(t_n, u_n, h)$ 的工作量太大, 一般不用 Taylor 展开法作数值计算, 但可用它计算附加初值.

4.2 单步法的稳定性和收敛性

将初值问题写成积分形式:

$$u(t+h) - u(t) = \int_t^{t+h} f(\tau, u(\tau))\mathrm{d}\tau, \quad u(t_0) = u_0. \tag{1.4.6}$$

如果有某一确定的函数 $\varphi(t, u, h)$ (通过某种离散化), 使初值问题的任一解 $u(t)$ 满足

$$u(t+h) - u(t) = h\varphi(t, u(t), h) + O(h^{p+1}), \tag{1.4.7}$$

其中 $p \geqslant 1$ 是使 (1.4.7) 成立的最大整数, 则称算法

$$u_{n+1} = u_n + h\varphi(t_n, u_n, h), \quad n = 0, 1, \cdots, T/h \tag{1.4.8}$$

为 p 阶单步法.

Taylor 展开法是 p 阶单步法, $\varphi(t, u, h)$ 由 (1.4.3) 定义. Euler 法是一阶单步法, 相应的 $\varphi = f(t, u(t))$.

注意

$$\varphi(t, u(t), h) = \frac{u(t+h) - u(t)}{h} + O(h^p) = \frac{1}{h}\int_t^{t+h} f(\tau, u(\tau))\mathrm{d}\tau + O(h^p),$$

故

$$\lim_{h \to 0} \varphi(t, u(t), h) = f(t, u(t)).$$

定义 $\varphi(t, u(t), 0) = f(t, u(t))$, 则知 $\varphi(t, u(t), h)$ 于 $h = 0$ 连续. 反之, 若 $\varphi(t, u(t), h)$ 于 $h = 0$ 连续且单步法 (1.4.8) 的局部截断误差 $R_h = o(h)(h \to 0)$, 则由

$$\varphi(t, u(t), h) = \frac{1}{h}\int_t^{t+h} f(\tau, u(\tau))\mathrm{d}\tau + o(1),$$

令 $h \to 0$, 知 $\varphi(t, u(t), 0) = f(t, u(t))$. 所以可将单步法的相容性定义为

定义 4.1 称单步法 (1.4.8) **相容**, 如果 $\varphi(t, u(t), h)$ 于 $h = 0$ 连续, 且

$$\varphi(t, u, 0) = f(t, u).$$

定理 4.1 设 $\varphi(t,u,h)(t_0 \leqslant t \leqslant T, 0 \leqslant h \leqslant h_0, u \in (-\infty,\infty))$ 关于 u 满足 Lipschitz 条件, 则单步法 (1.4.8) 稳定.

实际上, 设 v_n 是 (1.4.8) 的以 v_0 为初值的解, 则

$$v_{n+1} = v_n + h\varphi(t_n, v_n, h).$$

与 (1.4.8) 相减, 知 $e_n = u_n - v_n$ 满足

$$e_{n+1} = e_n + h[\varphi(t_n, u_n, h) - \varphi(t_n, v_n, h)],$$

于是

$$\begin{aligned} |e_{n+1}| &\leqslant |e_n| + hL|e_n| = (1+hL)|e_n| \\ &\leqslant \cdots \leqslant (1+hL)^{n+1}|e_0| \\ &\leqslant e^{L(n+1)h}|e_0| \leqslant e^{L(T-t_0)}|e_0|, (n+1)h \leqslant T - t_0, \end{aligned}$$

所以 (1.4.8) 稳定.

定理 4.2 设 $\varphi(t,u,h)$ 满足定理 4.1 的条件, 又单步法 (1.4.8) 相容, 则当 $h \to 0$ 时, 它的数值解 $u_n \to u(t)$, 只要 $t_0 + nh \to t$, 初值 $u_0 \to u(t_0)$. 若更设 (1.4.8) 是 p 阶单步法, 则还有敛速估计 (1.4.9).

证明 不妨设 (1.4.8) 是 p 阶单步法. 则初值问题的解 $u(t)$ 满足

$$u(t_{n+1}) = u(t_n) + h\varphi(t_n, u(t_n), h) + O(h^{p+1}).$$

与 (1.4.8) 相减, 并令 $e_n = u(t_n) - u_n$, 则

$$e_{n+1} = e_n + h[\varphi(t_n, u(t_n), h) - \varphi(t_n, u_n, h)] + O(h^{p+1}).$$

因 φ 关于 u 满足 Lipschitz 条件, 所以

$$\begin{aligned} |e_{n+1}| &\leqslant |e_n| + hL|e_n| + ch^{p+1}, \\ |e_{n+1}| - |e_n| &\leqslant hL|e_n| + ch^{p+1}. \end{aligned}$$

两端关于 n 求和, 得

$$|e_n| \leqslant (|e_0| + cTh^p) + hL\sum_{k=0}^{n-1}|e_k|.$$

再利用 Gronwall 不等式 (1.1.25), 就得到估计:

$$|e_n| \leqslant e^{LT}[(1+Lh)|e_0| + cTh^p]. \tag{1.4.9}$$

特别若取 $u_0 = u(t_0)$, 则 $e_0 = 0$, 误差 e_n 的阶为 $O(h^p)$.

4.3 Runge – Kutta 法

Taylor 展开法, 用 f 在同一点 (t_n, u_n) 的高阶导数表示 $\varphi(t_n, u_n, h)$, 这不便于数值计算. Runge – Kutta 型方法是用 f 在一些点上的值表示 $\varphi(t_n, u_n, h)$, 使单步法局部截断误差的阶和 Taylor 展开法相等. 我们先在区间 $[t, t+h]$ 上讨论. 将初值问题写成积分形式:

$$u(t+h) = u(t) + \int_t^{t+h} f(\tau, u(\tau))\mathrm{d}\tau. \tag{1.4.10}$$

在 $[t, t+h]$ 取 m 个点 $t_1 = t \leqslant t_2 \leqslant t_3 \leqslant \cdots \leqslant t_m \leqslant t+h$. 若知道 $k_i = f(t_i, u(t_i)), i = 1, \cdots, m$, 则可用它们的一次组合去近似 f:

$$\sum_{i=1}^{m} c_i k_i \approx f. \tag{1.4.11}$$

问题是如何计算 k_i (因 $u(t_i)$ 未知). 一个直观的想法是: 设已知 $(t_1, k_1) = (t_1, f(t_1, u(t_1)))$, 由 Euler 法 $u(t_2) \approx u(t_1)+(t_2-t_1)f(t_1, u(t_1)) = u(t_1)+(t_2-t_1)k_1$, 于是

$$k_2 \approx f(t_2, u(t_1) + (t_2 - t_1)k_1).$$

再利用 Euler 法又可以从 (t_2, k_2) 算出

$$k_3 \approx f(t_3, u(t_1) + (t_2 - t_1)k_1 + (t_3 - t_2)k_2).$$

如此可继续下去. 要求节点 $\{t_i\}$ 和系数 c_i 适当选取, 使近似式 (1.4.11) 有尽可能高的逼近阶. 这为下面构造 Runge–Kutta 法提供某些启示.

为便于推导, 我们先引进若干记号. 首先令

$$t_i = t + a_i h = t_1 + a_i h, \quad i = 2, \cdots, m,$$

其中 a_i 与 h 无关. 再引进下三角形系数阵:

$$\begin{matrix} b_{21} & & & \\ b_{31} & b_{32} & & \\ \vdots & \vdots & \ddots & \\ b_{m1} & b_{m2} & \cdots & b_{m,m-1}, \end{matrix}$$

其中 b_{ij} 与 h 无关,

$$\sum_{j=1}^{i-1} b_{ij} = a_i, \quad i = 2, 3, \cdots, m. \tag{1.4.12}$$

又 $c_i \geqslant 0$

$$\sum_{i=1}^{m} c_i = 1. \tag{1.4.13}$$

假设三组系数 $\{a_i\},\{b_{ij}\}$ 和 $\{c_i\}$ 已给定, 则 Runge–Kutta 法计算过程如下:

$$u_{n+1} = u_n + h\varphi(t_n, u_n, h), n = 0, 1, \cdots, \tag{1.4.14}$$

其中

$$\varphi(t, u(t), h) = \sum_{i=1}^{m} c_i k_i, \tag{1.4.15}$$

$$\begin{cases} k_1 = f(t, u), \\ k_2 = f(t + ha_2, u(t) + hb_{21}k_1), b_{21} = a_2, \\ k_3 = f(t + ha_3, u(t) + h(b_{31}k_1 + b_{32}k_2)), \quad b_{31} + b_{32} = a_3, \\ \cdots\cdots\cdots\cdots \\ k_m = f(t + ha_m, u(t) + h\sum\limits_{j=1}^{m-1} b_{mj}k_j), \quad \sum\limits_{j=1}^{m-1} b_{mj} = a_m. \end{cases} \tag{1.4.16}$$

系数 $\{a_i\},\{b_{ij}\}$ 和 $\{c_i\}$ 按如下原则确定: 将 k_i 关于 h 展开, 以之代到 (1.4.15), 使 l 次幂 $h^l(l = 0, 1, \cdots, p-1)$ 的系数和 (1.4.3) 同次幂的系数相等, 如此得到的算法 (1.4.14) 称为 m 级 p 阶 Runge–Kutta 法.

现在推导一些常用的计算方案. 将 $u(t+h)$ 展开到 h 的三次幂:

$$u(t+h) = u(t) + \sum_{l=1}^{3} \frac{h^l}{l!} u^{(l)}(t) + O(h^4) = u(t) + h\varphi_T(t, u, h), \tag{1.4.17}$$

其中

$$\begin{cases} \varphi_T(t, u, h) = f + \frac{1}{2}hF + \frac{1}{6}h^2(Ff_u + G) + O(h^3), \\ \qquad F = f_t + ff_u, \\ \qquad G = f_{tt} + 2ff_{tu} + f^2 f_{uu}. \end{cases} \tag{1.4.18}$$

其次, 由二元 Taylor 展式,

$$\begin{aligned} k_1 &= f(t, u) = f, \\ k_2 &= f(t + ha_2, u + ha_2k_1) \\ &= f + ha_2(f_t + k_1 f_u) + \frac{1}{2}h^2 a_2^2 (f_{tt} + 2k_1 f_{tu} + k_1^2 f_{uu}) + O(h^3) \\ &= f + ha_2F + \frac{1}{2}h^2 a_2^2 G + O(h^3). \end{aligned}$$

同样,

$$k_3 = f + ha_3F + h^2(a_2b_{32}f_uF + \frac{1}{2}a_3^2G) + O(h^3).$$

于是

$$\varphi(t,u,h) = (c_1+c_2+c_3)f + h(a_2c_2+a_3c_3)F + \frac{1}{2}h^2\left[2a_2b_{32}c_3f_uF + (a_2^2c_2+a_3^2c_3)G\right] + O(h^3). \tag{1.4.19}$$

比较 $\varphi(t,u,h)$ 和 $\varphi_T(t,u,h)$ 的同次幂系数, 可得以下具体方案.

(一) $m=1$ 比较 h 的零次幂, 知

$$\varphi(t,u,h) = f,$$

算法 (1.4.14) 是 Euler 法.

(二) $m=2$ 此时 $c_3=0$,

$$\varphi(t,u,h) = (c_1+c_2)f + ha_2c_2F + \frac{1}{2}h^2a_2^2c_2G + O(h^3).$$

与 $\varphi_T(t,u,h)$ 比较 $1,h$ 的系数, 则

$$c_1+c_2=1, \quad a_2c_2=\frac{1}{2}.$$

它有无穷多组解, 从而有无穷多个二级二阶算法. 两个常见的方法是

(1) $c_1=0,\ c_2=1,\ a_2=\frac{1}{2}$. 此时

$$u_{n+1} = u_n + hf\left(t_n+\frac{1}{2}h, u_n+\frac{1}{2}hf_n\right).$$

称为中点法, 这是一种修正的 Euler 法.

(2) $c_1=c_2=\frac{1}{2},\ a_2=1$, 此时

$$u_{n+1} = u_n + \frac{1}{2}h(f(t_n,u_n) + f(t_{n+1}, u_n+hf_n)).$$

这是改进的 Euler 法.

(三) $m=3$ 比较 (1.4.18) 和 (1.4.19), 令 $1,h,h^2$ 的系数相等, 并注意 F,G 的任意性, 得

$$c_1+c_2+c_3=1, \quad a_2c_2+a_3c_3=\frac{1}{2},$$
$$a_2^2c_2+a_3^2c_3=\frac{1}{3}, \quad a_2b_{32}c_3=\frac{1}{6}.$$

四个方程不能完全确定六个系数, 因此这是含两个参数的三级三阶方法. 常见的方案有

(1) **Heun 三阶方法**. 此时

$$c_1 = \frac{1}{4}, \quad c_2 = 0, \quad c_3 = \frac{3}{4},$$
$$a_2 = \frac{1}{3}, \quad a_3 = \frac{2}{3}, \quad b_{32} = \frac{2}{3}.$$

算法为

$$\begin{cases} u_{n+1} = u_n + \frac{h}{4}(k_1 + 3k_3), \\ k_1 = f(t_n, u_n), \\ k_2 = f\left(t_n + \frac{1}{3}h, u_n + \frac{1}{3}hk_1\right), \\ k_3 = f\left(t_n + \frac{2}{3}h, u_n + \frac{2}{3}hk_2\right). \end{cases} \tag{1.4.20}$$

(2) **Kutta 三阶方法**. 此时

$$c_1 = \frac{1}{6}, \quad c_2 = \frac{2}{3}, \quad c_3 = \frac{1}{6},$$
$$a_2 = \frac{1}{2}, \quad a_3 = 1, \quad b_{32} = 2.$$

算法为

$$\begin{cases} u_{n+1} = u_n + \frac{h}{6}(k_1 + 4k_2 + k_3), \\ k_1 = f(t_n, u_n), \\ k_2 = f\left(t_n + \frac{1}{2}h, u_n + \frac{1}{2}hk_1\right), \\ k_3 = f(t_n + h, u_n - hk_1 + 2hk_2). \end{cases} \tag{1.4.21}$$

当 f 与 u 无关时, 这就是 Simpson 公式.

(四) $m = 4$ 将 (1.4.18), (1.4.19) 展开到 h^3, 比较 $h^i (i = 0, 1, 2, 3)$ 的系数, 则得含 13 个待定系数的 11 个方程, 由此得到含两个参数的四级四阶 Runge–Kutta 方法类, 其中最常用的有以下两个算法:

四阶 Runge–Kutta 法:

$$\begin{cases} u_{n+1} = u_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4), \\ k_1 = f(t_n, u_n), \\ k_2 = f\left(t_n + \frac{1}{2}h, u_n + \frac{1}{2}hk_1\right), \\ k_3 = f\left(t_n + \frac{1}{2}h, u_n + \frac{1}{2}hk_2\right), \\ k_4 = f(t_n + h, u_n + hk_3). \end{cases} \tag{1.4.22}$$

和

$$
(1.4.23)\qquad \begin{cases} u_{n+1} = u_n + \dfrac{h}{8}(k_1 + 3k_2 + 3k_3 + k_4), \\ k_1 = f(t_n, u_n), \\ k_2 = f\left(t_n + \dfrac{1}{3}h, u_n + \dfrac{1}{3}hk_1\right), \\ k_3 = f\left(t_n + \dfrac{2}{3}h, u_n - \dfrac{1}{3}hk_1 + hk_2\right), \\ k_4 = f(t_n + h, u_n + hk_1 - hk_2 + hk_3). \end{cases}
$$

(1.4.22) 是最常用的 Runge–Kutta 法, 通常也称为经典的四阶 Runge–Kutta 法.

仿此人们还可造出更多的算法. 设 $f(t,u)(t_0 \leqslant t \leqslant T, u \in (-\infty, +\infty))$ 连续, 且关于 u 满足 Lipschitz 条件. 由 $\varphi(t, u(t), h)$ 的表达式 (1.4.15), (1.4.16) (注意 $c_1 + \cdots + c_m = 1$) 知 $\varphi(t, u, 0) = f(t, u)$, 即 Runge – Kutta 法相容, 所以定理 4.1 和定理 4.2 对 Runge – Kutta 法恒成立, 特别 p 阶 Runge – Kutta 法整体误差的阶为 $O(h^p)$.

例 4.1 用四级四阶 Runge – Kutta 法计算初值问题:

$$
\begin{aligned} u' &= 4tu^{\frac{1}{2}}, \quad 0 \leqslant t \leqslant 2, \\ u(0) &= 1. \end{aligned}
$$

取 $h = 0.1, 0.5, 1$. 精确解为

$$
u(t) = (1 + t^2)^2.
$$

计算结果如表 4.1. 与 §3 例 3.1 的表 3.1 比较, 可见 Runge–Kutta 法的稳定性较线性三阶二步法优越.

表 4.1

t	精确解	$h = 0.1$	$h = 0.5$	$h = 1$
0.0	1	1	1	1
0.3	1.081 600	1.081 599	–	–
0.5	1.562 500	1.562 497	1.561 106	–
0.8	2.689 600	2.689 592	–	–
1	4	3.999 985	3.993 247	3.913 900
1.3	5.953 600	5.953 576	–	–
1.5	10.562 500	10.562 455	10.542 656	–
2	25.000 000	24.999 904	24.957 954	24.530 977

习 题

就 $c_2 = c_3$ 和 $a_2 = a_3$ 导出三阶 Runge – Kutta 法.

§5 绝对稳定性和绝对稳定域

5.1 绝对稳定性

无论从理论或应用方面看, 单步和多步法都必须是稳定的. 但这种稳定有两个限制, 一是要求 $h \in (0, h_0)$ 充分小, 而实际用的 h 是固定的; 二是只允许初值有误差, 往后各步计算都精确, 而实际计算时每步都可能有舍入误差. 为了控制这种误差的增长, 需对多步法提出进一步要求, 即绝对稳定性.

对一般非线性常微分方程组:

$$\boldsymbol{u}' = \boldsymbol{f}(t, \boldsymbol{u})$$

讨论绝对稳定性是困难的, 通常是考虑它在解 $\boldsymbol{u}$ 临近的线性化方程:

$$(\boldsymbol{u} - \overline{\boldsymbol{u}})' = \frac{\partial f(t, \overline{\boldsymbol{u}})}{\partial \boldsymbol{u}}(\boldsymbol{u} - \overline{\boldsymbol{u}}).$$

为此我们讨论线性常微分方程组:

$$\boldsymbol{u}' = \boldsymbol{A}\boldsymbol{u}, \tag{1.5.1}$$

其中 $\boldsymbol{A}$ 是常数矩阵. 设 $\boldsymbol{A}$ 可对角化:

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \mathrm{diag}(\lambda_1, \cdots, \lambda_m),$$

λ_j 是实或复特征值, 作变换 $\boldsymbol{u} = \boldsymbol{T}\boldsymbol{v}$, 则 (1.5.1) 化为

$$v_i' = \lambda_i v_i, \quad i = 1, \cdots, m.$$

因此作为模型, 我们考虑方程:

$$u' = \mu u \tag{1.5.2}$$

的绝对稳定性, 其中 μ 是实数或复数.

求解 (1.5.2) 的线性多步法为

$$\sum_{j=0}^{k} \alpha_j u_{n+j} = \mu h \sum_{j=0}^{k} \beta_j u_{n+j}. \tag{1.5.3}$$

令 $\overline{h} = \mu h$, 则

$$\sum_{j=0}^{k} \alpha_j u_{n+j} - \overline{h} \sum_{j=0}^{k} \beta_j u_{n+j} = 0. \tag{1.5.4}$$

实际计算时每步都可能有舍入误差, 得到的是 u_n 的近似 $\overline{u}_n$, 满足

$$\sum_{j=0}^{k}(\alpha_j-\overline{h}\beta_j)\overline{u}_{n+j}=\eta_n, \tag{1.5.5}$$

其中 η_n 为局部舍入误差. 假设 $|\eta_n|\leqslant M$. 令 $\overline{e}_n=\overline{u}_n-u_n$, 显然 $\overline{e}_n$ 满足

$$\sum_{j=0}^{k}(\alpha_j-\overline{h}\beta_j)\overline{e}_{n+j}=\eta_n. \tag{1.5.6}$$

如前引进 k 维列向量 $\boldsymbol{E}_n=(\overline{e}_{n+k-1},\cdots,\overline{e}_n)^{\mathrm{T}}$ 和 $\boldsymbol{\eta}_n=(\alpha_k^{-1}\eta_n,0,\cdots,0)^{\mathrm{T}}$ 以及 $k\times k$ 矩阵

$$\overline{\boldsymbol{C}}=\begin{bmatrix}-a_k^{-1}a_{k-1} & -a_k^{-1}a_{k-2} & \cdots & -a_k^{-1}a_1 & -a_k^{-1}a_0\\ 1 & 0 & \cdots & 0 & 0\\ 0 & 1 & \cdots & 0 & 0\\ \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & 0\end{bmatrix}, \tag{1.5.7}$$

其中 $a_j=\alpha_j-\overline{h}\beta_j, j=k,k-1,\cdots,1,0$, 假定 h 充分小使 $a_k\neq 0$, 则可将 (1.5.6) 写成向量形式:

$$\boldsymbol{E}_n=\overline{\boldsymbol{C}}\boldsymbol{E}_{n-1}+\boldsymbol{\eta}_n.$$

逐次递推, 得

$$\boldsymbol{E}_n=\overline{\boldsymbol{C}}^n\boldsymbol{E}_0+(\overline{\boldsymbol{C}}^{n-1}\boldsymbol{\eta}_0+\cdots+\overline{\boldsymbol{C}}^{n-i-1}\boldsymbol{\eta}_i+\cdots+\boldsymbol{\eta}_{n-1}) \tag{1.5.8}$$

(参看 (1.3.17)). 为使误差 $\boldsymbol{E}_n$ 在逐步计算中减小, 应要求

$$\lim_{n\to\infty}\overline{\boldsymbol{C}}^n=\boldsymbol{O}. \tag{1.5.9}$$

令

$$\rho(\lambda)=\sum_{j=0}^{k}\alpha_j\lambda^j,\quad \sigma(\lambda)=\sum_{j=0}^{k}\beta_j\lambda^j, \tag{1.5.10}$$

则矩阵 $\overline{\boldsymbol{C}}$ 的特征方程为

$$\rho(\lambda)-\overline{h}\sigma(\lambda)=0. \tag{1.5.11}$$

由引理 1.1 的 (iii), 可知 (1.5.9) 成立的充要条件是 (1.5.11) 的根都在单位圆内.

定义 5.1 称线性 k 步法关于 $\overline{h}=\mu h$ **绝对稳定** (absolutely stable), 如果特征方程 (1.5.11) 的根都在单位圆内 ($|\lambda|<1$). 若存在复数域 D_A, 使多步法对 $\forall\overline{h}\in D_A$ 都绝对稳定, 则称 D_A 为绝对稳定域.

显然绝对稳定域越大, 方法的适应性越强, 因此也更优越.

可以证明 (参看 [2]), 当矩阵 $\overline{\boldsymbol{C}}$ 的特征值都在单位圆内时 ($\overline{\boldsymbol{C}}$ 的谱半径小于 1), 存在一种范数 $|||\cdot|||$, 使 $|||\overline{\boldsymbol{C}}|||<1$. 由 (1.5.8), 直接得

$$|||\boldsymbol{E}_n|||\leqslant|||\overline{\boldsymbol{C}}|||^n+M(1-|||\overline{\boldsymbol{C}}|||)^{-1},\quad M\geqslant|\eta_n|,\quad n=1,2,\cdots,$$

所以舍入误差的影响是可控的 (连续依赖 M).

对一般的非线性问题, f_u 不是常数, 此时可取 μ 是 f_u 的界或 f_u 的一个或多个代表值, h 为最大容许步长.

5.2 绝对稳定域

设 D_A 是多步法的绝对稳定域. 则特征方程 (1.5.11) 确定一由单位圆 $|\lambda|<1$ 到复平面 $\overline{h}\in Z$ 的解析变换:

$$\overline{h}=\rho(\lambda)/\sigma(\lambda). \tag{1.5.12}$$

命题 5.1 一个相容、稳定的多步法若绝对稳定, 则 μ 的实部 $\mathrm{Re}\mu<0$, 从而绝对稳定域 $D_A\subseteq Z_-$ (左半平面).

证明 因多步法相容且稳定, 故 1 必为 $\rho(\lambda)$ 的单根, 即 $\rho(1)=0$, 但 $\sigma(1)=\rho'(1)\neq0$. 设 k 次方程

$$\rho(\lambda)-\overline{h}\sigma(\lambda)=0$$

的根为 $\lambda_1(\overline{h}),\cdots,\lambda_k(\overline{h})$, 则必有唯一根, 比如 $\lambda_1(\overline{h})\to1$ (1 是 $\rho(\lambda)=0$ 的单根) 当 $\overline{h}\to0$, 其余根和 1 保持一正距离, 比如 $|\lambda_j(\overline{h})-1|\geqslant\delta_0>0$, 当 $\overline{h}\to0,j=2,\cdots,k$. 又 $u=\exp(\mu t)$ 满足 $u'=\mu u$, 若方法为 p 阶, 则

$$\sum_{j=0}^{k}\alpha_j\exp(\mu(t+jh))-\overline{h}\sum_{j=0}^{k}\beta_j\exp(\mu(t+jh))=O(h^{p+1})=O(\overline{h}^{p+1}),$$

从而

$$\rho(e^{\overline{h}})-\overline{h}\sigma(e^{\overline{h}})=O(\overline{h}^{p+1}).$$

将左端分解因子, 得

$$(e^{\overline{h}}-\lambda_1(\overline{h}))(e^{\overline{h}}-\lambda_2(\overline{h}))\cdots(e^{\overline{h}}-\lambda_k(\overline{h}))=O(\overline{h}^{p+1}).$$

因 $\exp(\overline{h}) \to 1$ 当 $\overline{h} \to 0$, 故除第一个因子外, 其余因子均有正下界, 即

$$\left|e^{\overline{h}} - \lambda_j(\overline{h})\right| \geqslant \delta_1 > 0, \quad j = 2, \cdots, k$$

(δ_1 是与 $\overline{h}$ 无关的正数), 因此

$$\lambda_1(\overline{h}) = e^{\overline{h}} + O(\overline{h}^{p+1}) = 1 + \overline{h} + O(\overline{h}^2) + O(\overline{h}^{p+1}). \tag{1.5.13}$$

由方法的绝对稳定性知 $|\lambda_1(\overline{h})| < 1$ 恒成立, 又 $p \geqslant 1$, 故必有 $\mathrm{Re}\overline{h} = \mathrm{Re}\mu h < 0$, 因此 $\mathrm{Re}\mu < 0$.

注意单位圆周 $|\lambda| = 1$ 在变换 (1.5.12) 之下, 所得的映像就是绝对稳定域 D_A 的边界. 由 (1.5.13), $\overline{h} = 0$ 时 $\lambda_1(0) = 1$, 故 D_A 的边界经过 0.

检验绝对稳定性归结为检验特征方程 (1.5.11) 的根是否在单位圆内 ($|\lambda| < 1$), 这有很多判别法, 如著名的 Routh–Hurwitz 准则,Schur 准则和 Miller 准则 (参看 [24] 的第 3 章). 这里只列一个简单而常用的判别法: 复系数二次方程 $\lambda^2 - b\lambda - c = 0$ 的根在单位圆内的充要条件是:

$$2 - |b|^2 - |b^2 + 4c| + 2|c|^2 > 0, \quad |c| < 1 \tag{1.5.14}$$

5.3 应用例子

例 5.1 Adams 法的绝对稳定域. Adams 外插法和内插法的特征多项式分别为

$$\rho(\lambda) - \overline{h}\sigma(\lambda) = \lambda^{k-1}(\lambda - 1) - \overline{h}\sum_{l=0}^{k-1} b_{k-1,l}\lambda^{k-l-1}$$

和

$$\rho(\lambda) - \overline{h}\sigma(\lambda) = \lambda^{k-1}(\lambda - 1) - \overline{h}\sum_{l=0}^{k} b_{kl}^*\lambda^{k-l},$$

其中 b_{kl} 和 b_{kl}^* 分别由表 2.2 及表 2.4 给出. 可以算出 k 步 $(k = 1, 2, 3, 4)$ Adams 外插法和内插法的绝对稳定域与实轴的交集分别为 $(\alpha_E, 0)$ 和 $(\alpha_I, 0)$, 其中 α_E, α_I 如表 5.1. 从表中看出, 内插法的稳定域比外插法大

表 5.1 Adams 方法的绝对稳定域

k	1	2	3	4
α_E	-2	-1	$-\dfrac{6}{11}$	$-\dfrac{3}{10}$
α_I	$-\infty$	-6	-3	$-\dfrac{90}{49}$

例 5.2 Runge – Kutta 法的绝对稳定域

在 Runge–Kutta 法中取 $f=\mu u$ (μ 为常数), 则

$$k_1=\mu u_n,$$
$$k_2=\mu(1+b_{21}\mu h)u_n=\mu P_1(\mu h)u_n,$$
$$k_3=\mu\left(u_n+h\sum_{j=1}^{2}b_{3j}k_j\right)=\mu(1+b_{31}\mu hP_1(\mu h))u_n=\mu P_2(\mu h)u_n,$$
$$\cdots\cdots\cdots\cdots$$
$$k_m=\mu P_{m-1}(\mu h)u_n,$$

其中 $P_i(\lambda)$ 是 i 次多项式, 从而

$$\varphi(t_n,u_n,h)=\mu\left(\sum_{i=1}^{m}c_iP_{i-1}(\mu h)\right)u_n.$$

Runge–Kutta 法为

$$u_{n+1}=u_n+h\varphi(t_n,u_n,h)=u_n+P_m(\mu h)u_n,\quad n=0,1,\cdots. \tag{1.5.15}$$

其次, 将 $u(t_{n+1})$ 在 t_n 展开:

$$u(t_{n+1})=u(t_n)+hu'(t_n)+\cdots+\frac{h^p}{p!}u^{(p)}(t_n)+O(h^{p+1}).$$

将 $u'=\mu u$ 的解 $u(t)=e^{\mu t}$ 代到上式得

$$u(t_{n+1})=\left(1+\mu h+\cdots+\frac{(\mu h)^p}{p!}\right)u(t_n)+O(h^{p+1}).$$

若方法是 p 阶的, 则

$$1+P_m(\mu h)=1+\mu h+\cdots+\frac{(\mu h)^p}{p!},$$

而 μh 任意, 故 $m\geqslant p$. 取 $m=p$, 将 (1,5.15) 写为

$$u_{n+1}=\left(\sum_{l=0}^{m}\frac{(\mu h)^l}{l!}\right)u_n,\quad n=0,1,\cdots,$$

其特征方程是一次的, 唯一的特征值

$$\lambda_1=1+\overline{h}+\frac{1}{2!}\overline{h}^2+\cdots+\frac{1}{m!}\overline{h}^m,\quad \overline{h}=\mu h. \tag{1.5.16}$$

注意, 当 $m=1,2,3,4$ 时, 解不等式 $|\lambda_1|<1$ 就可得出绝对稳定域 D_A, 参看图 5.1. 表 5.2 是 D_A 与实轴相交的区间, 从中看出, Runge–Kutta 法的绝对稳定域一般比线性多步法大, 这是它的优点.

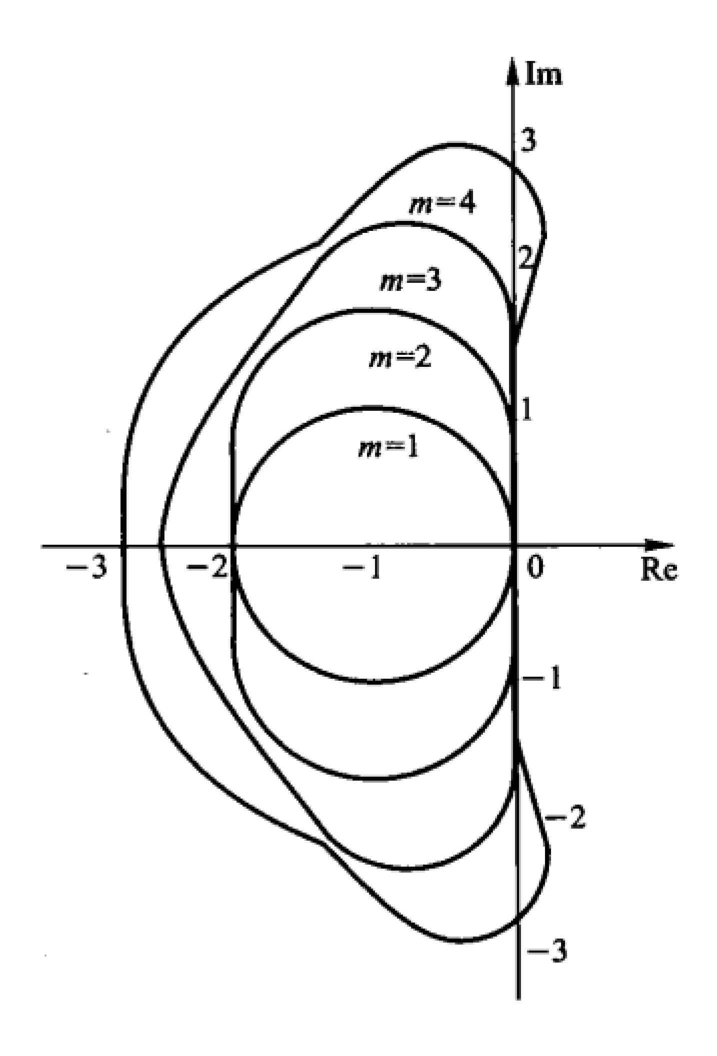


图 5.1

表 5.2

一级	λ_1	绝对稳定域
一级	$1+\overline{h}$	$(-2,0)$
二级	$1+\overline{h}+\dfrac{\overline{h}^2}{2}$	$(-2,0)$
三级	$1+\overline{h}+\dfrac{\overline{h}^2}{2}+\dfrac{\overline{h}^3}{6}$	$(-2.51,0)$
四级	$1+\overline{h}+\dfrac{\overline{h}^2}{2}+\dfrac{\overline{h}^3}{6}+\dfrac{\overline{h}^4}{24}$	$(-2.78,0)$

在使用 Runge–Kutta 法时, 应取 h 使 $\overline{h}$ 属于绝对稳定域, 否则有可能产生很大误差 (虽然方法稳定). 例如用四阶 Runge–Kutta 法 (1.4.22) 求解

$$u' = -20u, \quad u(0) = 1,$$

步长为 0.1 及 0.2. 当步长取为 0.1 时, $\overline{h}$ 属于绝对稳定域, 误差随 n 下降到零. 当步长取为 0.2 时, $\overline{h}$ 不属于绝对稳定域, 误差很快增长. 计算结果的误差如表 5.3.

表 5.3

t	$h=0.1$ 的误差	$h=0.2$ 的误差
0.0	0	0
0.2	−0.092 795	4.98
0.4	−0.012 010	25.0
0.6	−0.001 366	125.0
0.8	−0.000 152	625.0
1.0	−0.000 017	3 125.0

Runge–Kutta 法和预 – 校算法是求解常微分方程初值问题的有效数值方法. 本节介绍的 Runge–Kutta 法都是显式的, 为了进一步改善稳定域, 人们也采用隐式 Runge–Kutta 法, 参看 [22] — [24].

习　题

1. 证明方法 $u_{n+1}-u_n=hf_{n+1}$ 对所有 $\overline{h}\in(-\infty,0)$ 绝对稳定.
2. 验证表 5.1 所列 Adams 外插法和内插法的绝对稳定域.
3. 求二级二阶隐式 Runge–Kutta 法

$$\begin{aligned}&u_{n+1}=u_n+\frac{1}{2}h(k_1+k_2),\\&k_1=f(t_n,u_n),\\&k_2=f\left(t_n+h,u_n+\frac{1}{2}h(k_1+k_2)\right)\end{aligned}$$

的绝对稳定域.

§6　一阶方程组和刚性问题

6.1　对一阶方程组的推广

实际中遇到的不只是一阶方程式, 而且还有含 m 个方程的一阶方程组, 其一般形式为

$$\left\{\begin{aligned}&\frac{\mathrm{d}u_1}{\mathrm{d}t}=f_1(t,u_1,\cdots,u_m),\\&\frac{\mathrm{d}u_2}{\mathrm{d}t}=f_2(t,u_1,\cdots,u_m),\\&\cdots\cdots\cdots\cdots\\&\frac{\mathrm{d}u_m}{\mathrm{d}t}=f_m(t,u_1,\cdots,u_m),\end{aligned}\right.\tag{1.6.1}$$

这里 $f_1, \cdots, f_m(t_0 \leqslant t \leqslant T, |u_i| < \infty, i = 1, 2, \cdots, m)$ 是 $m+1$ 个变元的连续函数. 为使 (1.6.1) 的解确定, 还需给出初值条件:

(1.6.2) $$u_i(t_0) = u_{0i}, \quad i = 1, 2, \cdots, m.$$

这就是一阶方程组的初值问题. 如果实际问题不是一阶方程组而是高阶方程式, 我们也可以把它化为一阶方程组, 例如 m 阶微分方程:

$$u^{(m)} = f(t, u, u', \cdots, u^{(m-1)}),$$

只要引进新变量

$$u_1 = u, \quad u_2 = u', \cdots, \quad u_m = u^{(m-1)},$$

就化为一阶方程组:

$$\begin{aligned} &\frac{\mathrm{d}u_1}{\mathrm{d}t} = u_2, \\ &\frac{\mathrm{d}u_2}{\mathrm{d}t} = u_3, \\ &\cdots\cdots\cdots\cdots \\ &\frac{\mathrm{d}u_m}{\mathrm{d}t} = f(t, u_1, u_2, \cdots, u_{m-1}). \end{aligned}$$

此种转化不仅是理论上需要, 在计算上也可能更方便. 引进向量记号:

$$\boldsymbol{u} = (u_1, u_2, \cdots, u_m)^{\mathrm{T}}, \quad \boldsymbol{f} = (f_1, f_2, \cdots, f_m)^{\mathrm{T}}, \quad \boldsymbol{u}_0 = (u_{01}, u_{02}, \cdots, u_{0m})^{\mathrm{T}}$$

则 (1.6.1) 可写为向量形式:

(1.6.3) $$\boldsymbol{u}' = \boldsymbol{f}(t, \boldsymbol{u}), \quad \boldsymbol{u}(t_0) = \boldsymbol{u}_0.$$

若 $\boldsymbol{f}(t, \boldsymbol{u})$ 关于 $\boldsymbol{u}$ 满足 Lipschitz 条件, 则问题 (1.6.3) 有唯一解 (参看 [1]).

前面介绍的线性多步法, 预 – 校算法和 Runge–Kutta 法都可直接推广到一阶方程组, 只需用向量代替相应的标量. 所有关于相容性、稳定性和收敛性的定义和结论都可推广到方程组, 只需将绝对值号 $|\cdot|$ 换成 m 维欧氏空间的向量模 $\|\cdot\|$. 例如解方程组的线性多步法是:

(1.6.4) $$L[\boldsymbol{u}_n; h] = \sum_{j=0}^{k} [\alpha_j \boldsymbol{u}_{n+j} - h\beta_j \boldsymbol{f}(t_{n+j}, \boldsymbol{u}_{n+j})] = 0,$$

α_j, β_j 是标量, $L[\boldsymbol{u}_n; h]$ 是向量算子. 稳定性的充分必要条件仍然是第一特征多项式满足根条件.

但绝对稳定性的定义要适当修改. 因为在方程式的情形, f 的导数 $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}$ 是标量, 在模型问题 $\boldsymbol{u}' = \mu \boldsymbol{u}$ 中的 μ 也是标量. 现在 $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}$ 是 m 阶 Jacobi 矩阵, 应当用 m 阶常矩阵 $\boldsymbol{A}$ 代替 $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}$. 和 (1.5.6) 平行的线性化误差方程是:

$$\sum_{j=0}^{k} (\alpha_j \boldsymbol{I} - h\beta_j \boldsymbol{A}) \overline{\boldsymbol{e}}_{n+j} = \boldsymbol{\eta}_n. \tag{1.6.5}$$

假定矩阵 $\boldsymbol{A}$ 的特征值互异, 则可对角化

$$\boldsymbol{H}^{-1}\boldsymbol{A}\boldsymbol{H} = \begin{bmatrix} \mu_1 & & & 0 \\ & \mu_2 & & \\ & & \ddots & \\ 0 & & & \mu_m \end{bmatrix}$$

从而方程组 (1.6.5) 可化成 m 个独立的方程式, 相应的特征多项式有 m 个:

$$\rho(\lambda) - h\mu_l \sigma(\lambda) = 0, \quad l = 1, 2, \cdots, m, \tag{1.6.6}$$

其中

$$\rho(\lambda) = \sum_{j=0}^{k} \alpha_j \lambda^j, \sigma(\lambda) = \sum_{j=0}^{k} \beta_j \lambda^j.$$

以 $r_j^{(l)} (j = 1, 2, \cdots, k, \quad l = 1, 2, \cdots, m)$ 表示 (1.6.6) 的根的绝对值, 则可将方法的绝对稳定性定义为

$$r_j^{(l)} < 1, \quad j = 1, \cdots, k, \quad l = 1, 2, \cdots, m. \tag{1.6.7}$$

一般说来, 矩阵 $\boldsymbol{A}$ 非对称, 特征值 μ_l 是复的, 所以由 (1.6.7) 确定的参数 $\overline{h} = h\mu$ 的稳定域也是复的.

6.2 刚性问题

一阶方程组的解在 $t \to \infty$ 的过程中, 各分量的变化可能相差很大, 这会给计算带来一定困难. 例如考虑如下线性常微分方程组:

$$\boldsymbol{u}' = \boldsymbol{A}\boldsymbol{u}, \quad \boldsymbol{u} \in \mathbf{R}^3, \tag{1.6.8}$$

$$\boldsymbol{u}(0) = (2, 1, 2)^{\mathrm{T}}, \tag{1.6.9}$$

其中

$$\boldsymbol{A}=\begin{bmatrix}-0.1 & -49.9 & 0\\ 0 & -50 & 0\\ 0 & 70 & -3\times10^4\end{bmatrix}. \tag{1.6.10}$$

$\boldsymbol{A}$ 的特征值为

$$\lambda_1=-0.1,\quad \lambda_2=-50,\quad \lambda_3=-3\times10^4.$$

所以方程组 (1.6.8) – (1.6.9) 的解为

$$\begin{aligned}
&u_1(t)=\mathrm{e}^{-0.1t}+\mathrm{e}^{-50t},\\
&u_2(t)=\mathrm{e}^{-50t},\\
&u_3(t)=\frac{70}{3\times10^4-70}\mathrm{e}^{-50t}+\left(2-\frac{70}{3\times10^4-70}\right)\mathrm{e}^{-3\times10^{-4}}\\
&\qquad\ =0.002\,337\mathrm{e}^{-50t}+1.997\,663\mathrm{e}^{-3\times10^4t}.
\end{aligned}$$

显然当 $t\to\infty$ 时解的各分量 $u_i(t)$ 均按指数衰减到稳态解 $u_i(t)=0, i=1,2,3$. 但这些分量收敛到稳态解的速度很不一样, 这跟特征值的大小差别悬殊有关. 若用数值方法求解, 例如用 Runge – Kutta 法, 其绝对稳定域为 $(-\alpha_R,0)=(-2.78,0)$. 为使方法绝对稳定, 应要求 $|h\lambda_3|=3\times10^4h\leqslant\alpha_R=2.78$ 或

$$h\approx\frac{\alpha_R}{|\lambda_3|}=\frac{2.78}{3\times10^4}\approx10^{-4}.$$

其次, 为使 $u(t)$ 充分接近稳态解, 应要求 $\mathrm{e}^{\lambda_1t}=\mathrm{e}^{-0.1t}$ 充分接近零, 比如 $\mathrm{e}^{-0.1t}\approx\mathrm{e}^{-4}$ 或 $T\approx\dfrac{4}{|\lambda_1|}\approx40$, 于是计算步数

$$N\approx\frac{T}{h}\approx\frac{4|\lambda_3|}{\alpha_R|\lambda_1|}\approx\frac{40}{10^{-4}}=4\times10^5$$

若注意到每步要计算右端函数 f, 便知计算量是很大的.

产生上述困难主要来自方程的系数矩阵 (或一般方程的 Jacobi 矩阵) 的特征值的分布, 它们位于左半平面, 而按绝对值相差悬殊, 这类方程称为刚性方程 (Stiff 方程).

考虑线性常系数系统

$$\boldsymbol{u}'(t)=\boldsymbol{A}\boldsymbol{u}+\boldsymbol{g}(t) \tag{1.6.11}$$

和非线性系统

$$\boldsymbol{u}'(t)=\boldsymbol{f}(t,\boldsymbol{u}) \tag{1.6.12}$$

定义 6.1 称 (1.6.11) 为刚性的, 如果 λ_i 是矩阵 $\boldsymbol{A}$ 的特征值, 满足

(i) $\mathrm{Re}\lambda_i < 0, i = 1, 2, \cdots, m$,

(ii) $\max\limits_i |\mathrm{Re}\lambda_i| / \min\limits_i |\mathrm{Re}\lambda_i| = R >> 1$,

其中 $\mathrm{Re}\lambda_i$ 是 λ_i 的实部, R 称为刚性比.

定义 6.2 称 (1.6.12) 为刚性的, 如果在 t 的区间 $I = [0, T]$ 内, $\boldsymbol{f}$ 的 Jacobi 矩阵 $\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}$ 的特征值 $\lambda_i(t)$ 满足定义 6.1 的条件 (i)(ii).

粗略说来, 刚性方程组的特征是其解同时存在快变和慢变部分, 这类方程组在生物学、化学、电子学和控制论等领域有重要应用.

6.3 A 稳定性

刚性问题数值解法的基本问题是数值稳定性. 显然若刚性比 $R = \max\limits_i |\mathrm{Re}\lambda_i| / \min\limits_i |\mathrm{Re}\lambda_i|$ 很大, 则为保证方法绝对稳定, 最好要求绝对稳定域就是左复平面 Z_- : $\mathrm{Re}(h\lambda) < 0$, 为此引进 A 稳定定义.

定义 6.3 线性多步法说是 A 稳定的, 如果将它用于模型问题

$$u'(t) = \lambda u(t) \tag{1.6.13}$$

的绝对稳定域就是左复平面 Z_-, 其中 λ 是复数.

显然用 A 稳定算法计算时, 步长 h 不必受绝对稳定条件的限制.

为了判别线性多步法 A 稳定, 将它用于模型问题 (1.6.13), 得线性差分方程:

$$\sum_{j=0}^{k} (\alpha_j - \overline{h}\beta_j) u_{n+j} = 0, \quad \overline{h} = h\lambda \tag{1.6.14}$$

相应的特征方程为 $\rho(\lambda) - \overline{h}\sigma(\lambda) = 0$ 或

$$\overline{h} = \rho(\lambda)/\sigma(\lambda), \tag{1.6.15}$$

其中

$$\rho(\lambda) = \sum_{j=0}^{k} \alpha_j \lambda^j, \quad \sigma(\lambda) = \sum_{j=0}^{m} \beta_j \lambda^j, m \leqslant k. \tag{1.6.16}$$

(1.6.15) 是一个由 $\lambda \in Z$ 到 $\overline{h} \in Z$ 的单值解析变换, 其逆是由 $\overline{h} \in Z$ 到 $\lambda \in Z$ 的 k 值逆变换. 由定义推出

命题 6.1 设 $\lambda_i (i = 1, 2, \cdots, k)$ 是方程 (1.6.15) 的根, 则下列表述等价:

(i) 线性多步法 A 稳定;

(ii) $\mathrm{Re}\overline{h} < 0 \Rightarrow |\lambda_i| < 1, i = 1, 2, \cdots, k$;

(iii) $|\lambda| \geqslant 1 \Rightarrow \mathrm{Re}\overline{h}(\lambda) \geqslant 0$.

例 6.1 Euler 向后公式

$$u_{n+1} = u_n + hf_{n+1}.$$

由于 $\rho(\lambda) = \lambda - 1, \sigma(\lambda) = \lambda$, 则

$$\mathrm{Re}\overline{h}(\lambda) = \mathrm{Re}\frac{\lambda - 1}{\lambda} = \frac{|\lambda|^2 - |\lambda|\cos\theta}{|\lambda|^2} = \frac{|\lambda|(|\lambda| - \cos\theta)}{|\lambda|^2}.$$

显然当 $|\lambda| \geqslant 1$ 时, $\mathrm{Re}\overline{h}(\lambda) \geqslant 0$, 故 Euler 向后公式 A 稳定.

例 6.2 梯形公式

$$u_{n+1} = u_n + \frac{h}{2}(f_{n+1} + f_n).$$

因 $\rho(\lambda) = \lambda - 1, \sigma(\lambda) = \dfrac{1}{2}(\lambda + 1)$, 故

$$\mathrm{Re}\overline{h}(\lambda) = \mathrm{Re}\frac{\rho(\lambda)}{\sigma(\lambda)} = \mathrm{Re}\frac{\lambda - 1}{(\lambda + 1)/2} = 2\left(\frac{|\lambda|^2 - 1}{|\lambda + 1|^2}\right).$$

于是当 $|\lambda| \geqslant 1$ 时, $\mathrm{Re}\overline{h}(\lambda) \geqslant 0$. 所以梯形公式 A 稳定.

例 6.3 考虑 k 步线性方法:

$$u_{n+k} - u_n = \frac{h}{2}k(f_{n+k} + f_n).$$

因

$$\rho(\lambda) = \lambda^k - 1, \quad \sigma(\lambda) = \frac{k}{2}(\lambda^k + 1),$$

故

$$\overline{h} = \frac{\rho(\lambda)}{\sigma(\lambda)} = \frac{2}{k}\frac{\lambda^k - 1}{\lambda^k + 1} = \frac{2}{k}\frac{|\lambda|^{2k} - 1 + \mathrm{i}2|\lambda|^k\sin\theta}{|\lambda|^k + 1}.$$

于是

$$\mathrm{Re}\overline{h}(\lambda) = \frac{2}{k}\frac{|\lambda|^{2k} - 1}{|\lambda^k + 1|^2}.$$

显然当 $|\lambda| \geqslant 1$ 时, $\mathrm{Re}\overline{h}(\lambda) \geqslant 0$, 故此 k 步法 A 稳定.

现在看 Euler 法

$$u_{n+1} - u_n = hf_n.$$

这是显方法, 相应的 $\rho(\lambda) = \lambda - 1, \sigma(\lambda) = 1$, 此时

$$\mathrm{Re}\overline{h} = \mathrm{Re}\frac{1}{\lambda - 1} = \mathrm{Re}\frac{\mathrm{Re}\lambda - 1 + \mathrm{iIm}\lambda}{|\lambda - 1|^2} = \frac{\mathrm{Re}\lambda - 1}{|\lambda - 1|^2}.$$

显然若 $|\lambda| \geqslant 1$, 而 $\mathrm{Re}\lambda < 1$, 则 $\mathrm{Re}\overline{h} < 0$, 所以 Euler 法非 A 稳定.

实际上, 可以证明显线性多步法都不是 A 稳定的 (参看 [6]).

6.4　数值例子

1. 用向后 Euler 公式求解初值问题 (1.6.8)—(1.6.10). 以 u_j^n 表示 $u_j(t)$ 于 $t_n = nh$ 的近似, 则计算公式为

$$
\begin{aligned}
u_1^{n+1} - u_1^n &= h\left(-0.1u_1^{n+1} - 49.9u_2^{n+1}\right), \\
u_2^{n+1} - u_2^n &= h\left(-50u_2^{n+1}\right), \\
u_3^{n+1} - u_3^n &= h\left(70u_2^{n+1} - 3 \times 10^4 u_3^{n+1}\right)
\end{aligned}
$$

或

$$
\begin{aligned}
&(1 + 0.1h)\, u_1^{n+1} + 4.99hu_2^{n+1} = u_1^n, \\
&(1 + 50h)\, u_2^{n+1} = u_2^n, \\
&\left(1 + 3 \times 10^4 h\right) u_3^{n+1} - 70hu_2^{n+1} = u_3^n.
\end{aligned}
$$

以 $h = 1.0$ 和 $n = 1, 2, \cdots$ 计算, 直至

$$
\|u^n\| = [(u_1^n)^2 + (u_2^n)^2 + (u_3^n)^2]^{1/2} \leqslant e^{-4} \approx 0.000\ 027.
$$

计算结果如表 6.1. 从表中看到, $t = 100$ 时达到稳定解.

表 6.1

	精确解			数值解		
	$u_1(t)$	$u_2(t)$	$u_3(t)$	$u_1(t)$	$u_2(t)$	$u_3(t)$
$t = 0$	2.000 000	1.000 000	2.000 000	2.000 000	1.000 000	2.000 000
$t = 1$	0.904 837	0.000 000	0.000 000	0.928 698	0.019 607	0.000 112
$t = 2$	0.818 730	0.000 000	0.000 000	0.826 830	0.000 384	0.000 000
$t = 3$	0.740 818	0.000 000	0.000 000	0.751 322	0.000 007	0.000 000
$t = 4$	0.670 320	0.000 000	0.000 000	0.683 013	0.000 000	0.000 000
$t = 5$	0.606 530	0.000 000	0.000 000	0.620 921	0.000 000	0.000 000
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$t = 95$	0.000 082	0.000 000	0.000 000	0.000 128	0.000 000	0.000 000
$t = 96$	0.000 074	0.000 000	0.000 000	0.000 116	0.000 000	0.000 000
$t = 97$	0.000 067	0.000 000	0.000 000	0.000 106	0.000 000	0.000 000
$t = 98$	0.000 061	0.000 000	0.000 000	0.000 096	0.000 000	0.000 000
$t = 99$	0.000 055	0.000 000	0.000 000	0.000 087	0.000000	0.000 000
$t = 100$	0.000 050	0.000 000	0.000 000	0.000 087	0.000 000	0.000 000

2. 用向前 Euler 公式求解初值问题 (1.6.8)-(1.6.10), 计算公式为:

$$
\begin{aligned}
u_1^{n+1}-u_1^n &= -h(0.1hu_1^n+4.99u_2^n),\\
u_2^{n+1}-u_2^n &= -50hu_2^n,\\
u_3^{n+1}-u_3^n &= h\left(70u_2^n-3\times10^4u_3^n\right).
\end{aligned}
$$

取 $h=0,5$, 从表 6.2 看出, 计算到 $t=2$ 即出现不稳定现象.

表 6.2

	精确解			数值解		
	$u_1(t)$	$u_2(t)$	$u_3(t)$	$u_1(t)$	$u_2(t)$	$u_3(t)$
$t=0$	2.000 000	1.000 000	2.000 000	2.000 000	1.000 000	2.000 000
$t=0.5$	0.951 229	0.000 000	0.000 000	−23.050 000	−24.000 000	−29.963 000
$t=1$	0.904 837	0.000 000	0.000 000	576.902 499	576.000 000	$4.494\ 141\times10^8$
$t=1.5$	0.860 707	0.000 000	0.000 000	-13 823.142 625	−13 824.000 000	$-6.740\ 763\times10^{12}$
$t=2$	0.818 730	0.000 000	0.000 000	$3.317\ 768\times10^5$	$3.317\ 760\times10^5$	$1.011\ 047\times10^{17}$

*§7 外 推 法

7.1 多项式外推

设 $h>0$ 是离散化参数 (比如是步长), $A(h)$ 是按某种算法得到的数 A_0 的近似, 且 $\lim\limits_{h\to0}A(h)=A_0$ (定义 $A(0)=A_0$). 假定对任意正整数 $N, A(h)$ 有渐近展式:

$$
A(h)=A_0+A_1h+A_2h^2+\cdots+A_Nh^N+O(h^{N+1}),\quad h\to0 \tag{1.7.1}
$$

其中系数 $A_0,A_1,\cdots,A_N$ 与 h 无关. 若已算出 $A(h_0)$ 和 $A\left(\frac{1}{2}h_0\right)$, 则由 (1.7.1) 知

$$
2A\left(\frac{1}{2}h_0\right)-A(h_0)=A_0+O(h_0^2). \tag{1.7.2}
$$

可见左端较 $A(h_0)$ 和 $A\left(\frac{1}{2}h_0\right)$ 逼近 A_0 的阶更高. 这种由已知近似组合成更好近似的方法称为**外推法** (Extrapolation methods), 其基本思想属于 Richardson(1927). 计算数值积分的 Romberg 方法实际上也是一种外推法. 如果除 $A(h_0)$, $A\left(\frac{1}{2}h_0\right)$ 外还算出 $A\left(\frac{1}{4}h_0\right)$, 则可找到有逼近阶更高的线性组合:

$$
\frac{1}{3}A(h_0)-2A\left(\frac{h_0}{2}\right)+\frac{8}{3}A\left(\frac{h_0}{4}\right)=A_0+O(h_0^3). \tag{1.7.3}
$$

将这一思想推广，就导致更一般的所谓多项式外推法. 考虑一般的 h- 序列:

$$h_0 > h_1 > h_2 > \cdots > h_J > 0.$$

若展式 (1.7.1) 成立，则总可求得 $A(h_j)$ $(j=0,1,\cdots,J)$ 的线性组合，使

$$\sum_{j=0}^{J} c_{jJ} A(h_j) = A_0 + O(h_0^{J+1}). \tag{1.7.4}$$

实际上，求一以 $(h_j, A(h_j))$ $(j=0,1,\cdots,J)$ 为型值的 $A(h)$ 的 J 次插值多项式 $P_J(h)$，因

$$A(h) = P_J(h) + O(h_0^{J+1}),$$

$A(0) = A_0, P_J(0)$ 是 $A(h_j)$ 的一次组合，于上式令 $h=0$ 即得 (1.7.4). 求 $P_j(h)$ 可用 Aitken 的逐步线性插值实现. 例如以 $(h_0, a_0^{(0)})$ 和 $(h_1, a_1^{(0)})$ 为型值的 h 的一次多项式为

$$I_{01}(h) = \frac{1}{h_1 - h_0}\begin{vmatrix} a_0^{(0)} & h_0 - h \\ a_1^{(0)} & h_1 - h \end{vmatrix},$$

其中 $a_0^{(0)} = A(h_0), a_1^{(0)} = A(h_1)$. 令 $a_0^{(1)} = I_{01}(0)$，则由 (1.7.4) 知 $a_0^{(1)} = A_0 + O(h_0^2)$. 特别当 $h_1 = \frac{1}{2}h_0$ 时，$a_0^{(1)}$ 就是 (1.7.2) 的左端. 同样过型值 $(h_1, a_1^{(0)})$, $(h_2, a_2^{(0)})(a_2^{(0)} = A(h_2))$ 的一次多项式为

$$I_{12}(h) = \frac{1}{h_2 - h_1}\begin{vmatrix} a_1^{(0)} & h_1 - h \\ a_2^{(0)} & h_2 - h \end{vmatrix}.$$

令 $a_1^{(1)} = I_{12}(0)$. 以 $I_{012}(h)$ 表示过三个点 $(h_0, a_0^{(0)}), (h_1, a_1^{(0)}), (h_2, a_2^{(0)})$ 的二次多项式，则

$$I_{012}(h) = \frac{1}{h_2 - h_0}\begin{vmatrix} I_{01}(h) & h_0 - h \\ I_{12}(h) & h_2 - h \end{vmatrix}.$$

令 $a_0^{(2)} = I_{012}(0)$，则由 (1.7.4) 知 $a_0^{(2)} = A_0 + O(h_0^3)$. 特别当 $h_1 = \frac{h_0}{2}, h_2 = \frac{h_0}{4}$，则 $a_0^{(2)}$ 就是 (1.7.3) 的左端. 如此可继续做下去，但插值多项式的次数不能超过方法的阶. 实际应用时，作一两次插值就够了.

若渐近展式形如

$$A(h) = A_0 + A_2 h^2 + A_4 h^4 + O(h^6),$$

则线性组合

$$\frac{4}{3}A(\frac{1}{2}h_0) - \frac{1}{3}A(h_0) = A_0 + O(h_0^4), \tag{1.7.5}$$

比 (1.7.2) 的收敛阶更高，可见外推法的效果和渐近展开的类型有关.

7.2 对初值问题的应用

设 $u(t)$ 是初值问题

$$u' = f(t, u), \quad u(t_0) = u_0$$

的解, $u(t, h)$ 是由某一数值方法 (如线性多步法, 预 - 校算法,Runge-Kutta 法等) 确定的 $u(t)$ 的近似解, h 是步长. 假定 H 是基本步长 (可取大一些), $h_0 = H/N_0(N_0 \geqslant 1)$. 从 $t = 0$ 出发, 用数值方法算 N_0 步, 得到 $u(t_0 + H, h_0)$. 然后取 $h_1 = H/N_1, N_1 > N_0$, 从 $t = t_0$ 出发用数值方法算 N_1 步, 得到近似解 $u(t_0 + H; h_1)$. 一般取 $h_j = H/N_j, j = 0, 1, \cdots, J, N_0 < N_1 < \cdots < N_J$, 从 $t = t_0$ 出发用数值方法算 N_j 步, 得到近似解 $u(t_0 + H, h_j)$. 如果数值解有渐近展开:

$$u(t, h) = u(t) + A_1 h + A_2 h^2 + A_3 h^3 + \cdots + A_N h^N + O(h^{N+1}), \quad h \to 0. \tag{1.7.6}$$

令 $a_j^{(0)} = u(t_0 + H, h_j), j = 0, 1, \cdots, J$, 则可按逐步线性插值法得到 $u(t)$ 的更为精确的近似. 通常取 $h_0 = H, h_1 = \dfrac{1}{2}h_0$, 则

$$u(H, h_0) = u(H) + O(h_0),$$

$$2u(H, h_1) - u(H, h_0) = u(H) + O(h_0^2). \tag{1.7.7}$$

对单步法, 常有展式:

$$u(t, h) = u(t) + \varepsilon(t) h^p + O(h^{p+1}), \tag{1.7.8}$$

其中 $\varepsilon(t)$ 为误差主项系数. 用 $\dfrac{1}{2}h$ 为步长由 t_0 算到 t, 则

$$u\left(t, \frac{1}{2}h\right) = u(t) + \varepsilon(t)\left(\frac{h}{2}\right)^p + O(h^{p+1}). \tag{1.7.9}$$

与 (1.7.8) 作线性组合 (线性插值), 得

$$\frac{2^p u\left(t, \dfrac{h}{2}\right) - u(t, h)}{2^p - 1} = u(t) + O(h^{p+1}). \tag{1.7.10}$$

7.3 用外推法估计误差

假定方法的数值解 $u(t, h)$ 有渐进展式 (1.7.8). 以 $e(t, h) = u(t) - u(t, h)$ 表示整体误差. 由 (1.7.8), (1.7.9) 得

$$e\left(t, \frac{h}{2}\right) = 2^{-p} e(t, h) + O(h^{p+1}).$$

而 $e\left(t,\frac{h}{2}\right)=u(t)-u\left(t,\frac{h}{2}\right)=e(t,h)+u(t,h)-u\left(t,\frac{h}{2}\right)$, 代入上式左端, 解出

$$e(t,h)=\frac{2^p}{2^p-1}\left[u\left(t,\frac{h}{2}\right)-u(t,h)\right]+O(h^{p+1}). \tag{1.7.11}$$

取

$$\overline{e}(t,h)=\frac{2^p}{2^p-1}\left[u\left(t,\frac{h}{2}\right)-u(t,h)\right] \tag{1.7.12}$$

作为 $e(t,h)$ 的近似, 右端是可以用近似解估算的. 利用 (1.7.12) 可以估计解的误差.

例 用 Euler 法解初值问题:

$$u'=-u,\quad u(0)=1.$$

其精确解 $u(t)=\mathrm{e}^{-t}$. Euler 法的误差阶为 $O(h)$, 用外推公式 (1.7.7) 的误差阶为 $O(h^2)$. 表 7.1 就 $h=2^{-j}, j=0,1,2,3,4,5,6$ 列出 $t=1$ 的近似值和误差. 从表 7.1 中看出, 外推解比 Euler 解精确, 用 (1.7.12) 作出的误差估计与实际误差基本符合.

表 7.1

h	Euler 解 $u(1,h)$	$2u\left(1,\frac{1}{2}\right)-u(1,h)$	Euler 解的误差	外推解误差	用 (1.7.12) 估计出的误差
1	0.000 000	0.500 000	−0.367 879	0.132 121	−0.500 000
$\frac{1}{2}$	0.250 000	0.382 813	−0.117 879	0.014 933	−0.132 813
$\frac{1}{4}$	0.316 406	0.370 812	−0.051 473	0.002 932	−0.054 405
$\frac{1}{8}$	0.343 609	0.368 539	−0.024 271	0.000 660	−0.024 930
$\frac{1}{16}$	0.356 074	0.368 036	−0.011 805	0.000 157	−0.011 962
$\frac{1}{32}$	0.362 055	0.367 918	−0.005 824	0.000 038	−0.005 863

习　　题

1. 证明任一相容的显式单步法的解 $u(t,h)$ 有渐近展式 (1.7.8).
2. 设单位圆内正 n 边型的半周长为 $A(h), nh=1$, 证明 $A(h)$ 可展成:

$$A(h)=\pi+A_2h^2+A_4h^4+\cdots.$$

并利用 $A(h), h=\frac{1}{4},\frac{1}{6},\frac{1}{8}$ 推出 π 的近似式.

第二章　椭圆型方程的有限差分法

从本章开始，我们介绍偏微分方程的数值解法，主要是有限差分法和 Galerkin 有限元法. 由于计算机只能存储有限个数据和做有限次运算, 所以任何一种用计算机解题的方法, 都必须把连续问题 (微分方程的边值问题、初值问题等) 离散化, 最终化成有限形式的线性代数方程组. 用差分法和有限元法将连续问题离散化的步骤是, 首先对求解区域做网格剖分, 用有限个网格节点代替连续区域, 然后将微分算子离散化, 从而把微分方程的定解问题化为线性代数方程组的求解问题. 差分法和有限元法的主要区别是离散化的第二步. 前者从定解问题的微分形式或积分形式出发, 用数值微商或数值积分公式导出相应的线性代数方程组. 后者从定解问题的变分形式出发, 用 Ritz-Galerkin 法导出相应的线性代数方程组, 但基函数按特定方式选取. 本章和第三章、第四章讨论有限差分法, 第五章、第六章讨论 Galerkin 有限元法.

差分法的基本问题有:

(1) 对求解域作网格剖分.

一维情形是把区间分成一些等距或不等距的小区间, 称之为**单元**. 二维情形则把区域分割成一些均匀或不均匀的矩形, 其边与坐标轴平行. 也可分割成一些小三角形或凸四边形等.

(2) 构造逼近微分方程定解问题的差分格式.

我们将介绍两种构造差分格式的方法: 直接差分化法和有限体积法. 也就两点边值问题介绍待定系数法的基本思想.

(3) 差分解的存在唯一性、收敛性及稳定性的研究.

(4) 差分方程的解法, 包括逐次超松弛法、交替方向法和共轭梯度法.

§1　差分逼近的基本概念

考虑二阶常微分方程边值问题:

$$Lu = -\frac{\mathrm{d}^2u}{\mathrm{d}x^2} + qu = f, \quad a < x < b, \tag{2.1.1}$$

$$u(a) = \alpha, \quad u(b) = \beta, \tag{2.1.2}$$

其中 q, f 为 $[a, b]$ 上的连续函数, $q \geqslant 0; \alpha, \beta$ 为给定常数. 这是最简单的椭圆方程第一边值问题.

将区间 $[a, b]$ 分成 N 等分, 分点为

$$x_i = a + ih \quad i = 0, 1, \cdots, N,$$

$h = (b-a)/N$. 于是我们得到区间 $I = [a, b]$ 的一个网格剖分. x_i 称为网格的**节点**, h 称为**步长**.

现在将方程 (2.1.1) 在节点 x_i 离散化. 为此, 对充分光滑的解 u, 由 Taylor 展式可得

$$\begin{aligned}\frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} \\ = \left[\frac{\mathrm{d}^2 u(x)}{\mathrm{d}x^2}\right]_i + \frac{h^2}{12}\left[\frac{\mathrm{d}^4 u(x)}{\mathrm{d}x^4}\right]_i + O(h^3),\end{aligned} \tag{2.1.3}$$

其中 $[\ ]_i$ 表示方括号内的函数在 x_i 点取值. 于是在 x_i 可将方程 (2.1.1) 写成

$$-\frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} + q(x_i)u(x_i) = f(x_i) + R_i(u), \tag{2.1.4}$$

其中

$$R_i(u) = -\frac{h^2}{12}\left[\frac{\mathrm{d}^4 u(x)}{\mathrm{d}x^4}\right]_i + O(h^3). \tag{2.1.5}$$

显然, 当 $h \to 0$ 时, $R_i(u)$ 是 h 的二阶无穷小量. 若舍去 $R_i(u)$, 则得到逼近方程 (2.1.1) 的差分方程:

$$L_h u_i = -\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + q_i u_i = f_i, \tag{2.1.6}$$

式中 $q_i = q(x_i), f_i = f(x_i)$. 称 $R_i(u)$ 为差分方程 (2.1.6) 的**截断误差**. 利用差分算子 L_h, 可将 (2.1.4) 写成

$$L_h u(x_i) = f(x_i) + R_i(u). \tag{2.1.4$'$}$$

而在节点 x_i 处, 微分方程 (2.1.1) 为

$$[Lu]_i = f(x_i),$$

与 (2.1.4)$'$ 相减, 得

$$R_i(u) = L_h u(x_i) - [Lu]_i. \tag{2.1.7}$$

所以 $R_i(u)$ 是用差分算子 L_h 逼近微分算子 L 所引起的截断误差, 在这里关于 h 的阶为 $O(h^2)$.

差分方程 (2.1.6) 当 $i=1,2,\cdots,N-1$ 时成立, 加上边值条件 $u_0=\alpha, u_N=\beta$, 就得到关于 u_i 的线性代数方程组:

$$L_h u_i = -\frac{u_{i+1}-2u_i+u_{i-1}}{h^2}+q_iu_i=f_i, \quad i=1,2,\cdots,N-1, \tag{2.1.8}$$

$$u_0=\alpha, \quad u_N=\beta. \tag{2.1.9}$$

它的解 u_i 是 $u(x)$ 于 $x=x_i$ 的近似. 称 (2.1.8)-(2.1.9) 为逼近 (2.1.1)-(2.1.2) 的**差分方程**或**差分格式**. 由于 (2.1.8) 是用二阶中心差商代替 (2.1.1) 中二阶微商得到的, 所以也称 (2.1.8)-(2.1.9) 为**中心差分格式**.

注意方程 (2.1.8) 的个数等于网格内点 $x_1,x_2,\cdots,x_{N-1}$ 的个数 $N-1$, 因此它是 $N-1$ 阶的方程组. 一般说来, 这是高阶方程组, 例如取 $N=100$ (即把区间 $[a,b]$ 作 100 等分), 则阶数为 99. 但每个方程的未知数最多有三个, 因此系数矩阵 $\boldsymbol{A}$ 的大量元素是零. 如果我们把方程和未知数由左到右排列, 则 $\boldsymbol{A}$ 是对称的三对角矩阵. 例如取 $N=5$, 则

$$\boldsymbol{A}=\begin{bmatrix} \frac{2}{h^2}+q_1 & -\frac{1}{h^2} & 0 & 0 \\ -\frac{1}{h^2} & \frac{2}{h^2}+q_2 & -\frac{1}{h^2} & 0 \\ 0 & -\frac{1}{h^2} & \frac{2}{h^2}+q_3 & -\frac{1}{h^2} \\ 0 & 0 & -\frac{1}{h^2} & \frac{2}{h^2}+q_4 \end{bmatrix}.$$

我们可用消元法或迭代法求解方程组 (2.1.8)-(2.1.9).

对于差分方程 (2.1.8)-(2.1.9), 我们自然关心它是否有唯一解. 其次, 差分解 u_i 当网格无限加密, 或者说 $h\to 0$ 时是否收敛到精确解 $u(x_i)$, 以及在何种度量意义下收敛, 收敛的速度如何. 为此需要引进若干记号.

以 I_h 表示网格内点 $x_1,x_2,\cdots,x_{N-1}$ 的集合, $\overline{I}_h$ 表示网格内点和界点 $x_0=a$, $x_N=b$ 的集合. 定义在 I_h (相应的 $\overline{I}_h$) 上的函数 $u_h(x_i)=u_i$ 称为 I_h (相应的 $\overline{I}_h$) 上的**网格函数**. 和连续变量的函数类似, 我们对 I_h 上的网格函数引进范数

$$\|u_h\|_C=\max_{1\leqslant i\leqslant N-1}|u_i|, \tag{2.1.10}$$

$$\|u_h\|_0^2=\sum_{i=1}^{N-1}hu_i^2. \tag{2.1.11}$$

$$\|u_h\|_1^2=\|u_h\|_0^2+|u_h|_1^2, \tag{2.1.12}$$

其中

$$|u_h|_1^2=\sum_{i=1}^{N}h\left(\frac{u_i-u_{i-1}}{h}\right)^2. \tag{2.1.13}$$

若不特别说明, 我们用 $\|\cdot\|$ 表示 (2.1.10)~(2.1.12) 中任一种范数.

很明显, 要想差分解 u_h 按范数 $\|\cdot\|$ 收敛到 u, 差分算子 L_h 必须在某种意义下逼近微分算子 L, 这导致下列定义.

定义 1.1 设 $\mathcal{M}$ 是某一充分光滑的函数类, $R_h(u)$ 是由截断误差 (2.1.7) 定义的网格函数. 若对任何 $u\in\mathcal{M}$, 恒有

$$\lim_{h\to 0}\|R_h(u)\|=0, \tag{2.1.14}$$

则说差分算子 L_h 逼近微分算子 L, 而称 (2.1.14) 为**相容条件**.

由 (2.1.5) 便知, 差分算子 (2.1.6) 逼近微分算子 (2.1.1), 且逼近的阶是: $\|R_h(u)\|_C=O(h^2), \|R_h(u)\|_0=O(h^2), \|R_h(u)\|_1=O(h)$.

定义 1.2 称差分解 u_h 收敛到边值问题的解 u, 如果当 h 充分小后, (2.1.8), (2.1.9) 的解 u_h 存在, 且按某一范数 $\|\cdot\|$ 有

$$\lim_{h\to 0}\|u_h-u\|=0, \tag{2.1.15}$$

这里把 u 看成 I_h 上的网格函数.

将 (2.1.4) 写成

$$L_hu(x_i)=f_i+R_i(u),$$

以此与 (2.1.8) 相减, 得

$$L_h(u(x_i)-u_i)=R_i(u).$$

引进误差

$$e_i=u(x_i)-u_i,$$

则误差函数 $e_h(x_i)=e_i$ 满足下列差分方程:

$$\begin{cases}L_he_i=R_i(u), & i=1,2,\cdots,N-1.\\ e_0=e_N=0,\end{cases} \tag{2.1.16}$$

于是收敛性及收敛速度的估计问题, 就归结到通过右端 $R_i(u)$ (截断误差) 估计误差函数 e_h 的问题. 这和差分方程的稳定性有关.

定义 1.3 称差分方程 $L_h v_i = f_i \ (i = 1, 2, \cdots, N-1)$, $v_0 = v_N = 0$ 关于右端稳定, 如果存在与网格 I_h 及右端 $f_h(f_h(x_i) = f_i)$ 无关的正常数 M 和 h_0, 使

$$(2.1.17) \qquad \|v_h\| \leqslant M\|f_h\|_R, \quad 当\ 0 < h < h_0,$$

其中 $\|f_h\|_R$ 是右端 f_h 的某一范数, 它可以和 $\|\cdot\|$ 相同, 也可以不同, $v_h(x_i) = v_i$, $i = 1, 2, \cdots, N-1$.

不等式 (2.1.17) 表明, 解 v_h 连续依赖右端 f_h, 即右端变化小时解的变化也小. 实际上, 设 $u_h^{(1)}, u_h^{(2)}$ 是差分方程 (2.1.8), (2.1.9) 对应右端 $f_h^{(1)}, f_h^{(2)}$ 的解, 则 $v_h = u_h^{(1)} - u_h^{(2)}$ 满足 $L_h v_i = f_i^{(1)} - f_i^{(2)}$, $v_0 = v_N = 0$. 由 (2.1.17),

$$\|u_h^{(1)} - u_h^{(2)}\| \leqslant M\|f_h^{(1)} - f_h^{(2)}\|_R.$$

若 $\lim \|f_h^{(1)} - f_h^{(2)}\|_R = 0$, 则 $\lim \|u_h^{(1)} - u_h^{(2)}\| = 0$.

由 (2.1.17) 推出, 与 (2.1.8), (2.1.9) 相应的齐方程 ($f_i = 0, \alpha = \beta = 0$) 只能有平凡解 $u_i \equiv 0 \ (i = 1, 2, \cdots, N-1)$, 从而非齐方程对任何边值及右端有唯一解.

将不等式 (2.1.17) 用到误差方程 (2.1.16), 则

$$(2.1.18) \qquad \|e_h\| \leqslant M\|R_h(u)\|_R.$$

若解 u 充分光滑, L_h 关于范数 $\|\cdot\|$ 满足相容条件, 则当 $h \to 0$ 时 $\|e_h\| \to 0$, 从而差分解收敛到边值问题的解, 且有和截断误差相同的收敛阶.

定理 1.1 若边值问题的解 u 充分光滑, 差分方程按 $\|\cdot\|_R$ 满足相容条件, 且关于右端稳定, 则差分解 u_h 按 $\|\cdot\|$ 收敛到边值问题的解, 且有和 $\|R_h(u)\|_R$ 相同的收敛阶.

这样, 为了建立差分解的收敛性, 就需要检验相容条件和建立差分方程的稳定性. 检验相容条件并不困难, 例如由 (2.1.8) 定义的差分算子, 我们曾用 Taylor 展式证明它关于 $\|\cdot\|_0$ 及 $\|\cdot\|_C$ 都满足相容条件, 并且估计了截断误差的阶. 我们的主要问题是去建立差分方程的稳定性, 即建立形如 (2.1.17) 的估计式, 称之为关于差分方程解的**先验估计**.

稳定性概念在理论研究和实际应用中都有重要意义. 实际上, 由于有实测误差和舍入误差, 右端数据不可能准确给出. 如果小的右端误差会引起解的很大偏离, 即差分方程不稳定, 便失去实际意义.

§2 一维差分格式

考虑两点边值问题:

$$Lu = -\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}u}{\mathrm{d}x}\right) + r\frac{\mathrm{d}u}{\mathrm{d}x} + qu = f, \quad a < x < b, \tag{2.2.1}$$

$$u(a) = \alpha, \quad u(b) = \beta. \tag{2.2.2}$$

假定 $p \in C^1[a,b], p(x) \geqslant p_{\min} > 0, r, q, f \in C[a,b], \alpha, \beta$ 是给定的常数.

本节我们将介绍构造差分格式的三种方法: 直接差分化法、有限体积法和待定系数法. 还将讨论边值条件的逼近方法.

显然, 我们可以造出许多逼近 (2.2.1)–(2.2.2) 的差分格式, 但并非任何格式都是可取的. 一个好的差分格式, 应该是以尽可能小的工作量 (包括程序的准备和计算机的运算) 获得所需精度的结果. 因此, 一方面, 差分格式应该结构简单、便于求解; 另一方面, 应具有尽可能高的逼近阶. 此外, 还要根据问题的特点, 对差分格式提出其他要求.

2.1 直接差分化

首先取 $N+1$ 个节点:

$$a = x_0 < x_1 < \cdots < x_i < \cdots < x_N = b,$$

将区间 $I = [a,b]$ 分成 N 个小区间:

$$I_i : x_{i-1} \leqslant x \leqslant x_i, \quad i = 1, 2, \cdots, N.$$

于是得到 I 的一个**网格剖分**. 记 $h_i = x_i - x_{i-1}$, 称 $h = \max\limits_i h_i$ 为最大网格步长. 用 I_h 表示网格内点 $x_1, x_2, \cdots, x_{N-1}$ 的集合, $\overline{I}_h$ 表示内点和界点 $x_0 = a, x_N = b$ 的集合.

取相邻节点 x_{i-1}, x_i 的中点 $x_{i-\frac{1}{2}} = \dfrac{1}{2}(x_{i-1} + x_i)$ $(i = 1, 2, \cdots, N)$, 称为半整数点. 则由节点

$$a = x_0 < x_{\frac{1}{2}} < x_{\frac{3}{2}} < \cdots < x_{i-\frac{1}{2}} < \cdots < x_{N-\frac{1}{2}} < x_N = b$$

又构成 $[a,b]$ 的一个剖分, 称为**对偶剖分**. 图 2.1 中打 "•" 号的是原剖分节点, 打 "×" 号的是对偶剖分节点.

其次用差商代替微商, 将方程 (2.2.1) 在内点 x_i 离散化. 注意对充分光滑的

a x_{i-1} $x_{i-\frac{1}{2}}$ x_i $x_{i+\frac{1}{2}}$ x_{i+1} b

图 2.1

u, 由 Taylor 展式有

(2.2.3)
$$\frac{u(x_{i+1})-u(x_{i-1})}{h_i+h_{i+1}}=\left[\frac{\mathrm{d}u}{\mathrm{d}x}\right]_i+\frac{h_{i+1}-h_i}{2}\left[\frac{\mathrm{d}^2u}{\mathrm{d}x^2}\right]_i+O(h^2),$$

(2.2.4)
$$\begin{aligned}p\left(x_{i-\frac{1}{2}}\right)\frac{u(x_i)-u(x_{i-1})}{h_i}&=\left[p\frac{\mathrm{d}u}{\mathrm{d}x}\right]_{i-\frac{1}{2}}+\frac{h_i^2}{24}\left[p\frac{\mathrm{d}^3u}{\mathrm{d}x^3}\right]_{i-\frac{1}{2}}+O(h^3)\\&=\left[p\frac{\mathrm{d}u}{\mathrm{d}x}\right]_{i-\frac{1}{2}}+\frac{h_i^2}{24}\left[p\frac{\mathrm{d}^3u}{\mathrm{d}x^3}\right]_i+O(h^3),\end{aligned}$$

(2.2.5)
$$p\left(x_{i+\frac{1}{2}}\right)\frac{u(x_{i+1})-u(x_i)}{h_{i+1}}=\left[p\frac{\mathrm{d}u}{\mathrm{d}x}\right]_{i+\frac{1}{2}}+\frac{h_{i+1}^2}{24}\left[p\frac{\mathrm{d}^3u}{\mathrm{d}x^3}\right]_i+O(h^3).$$

由 (2.2.5) 减 (2.2.4), 并除以 $\dfrac{(h_i+h_{i+1})}{2}$, 得

(2.2.6)
$$\begin{aligned}&\frac{2}{h_i+h_{i+1}}\left[p\left(x_{i+\frac{1}{2}}\right)\frac{u(x_{i+1})-u(x_i)}{h_{i+1}}-p\left(x_{i-\frac{1}{2}}\right)\frac{u(x_i)-u(x_{i-1})}{h_i}\right]\\&=\frac{2}{h_i+h_{i+1}}\left(\left[p\frac{\mathrm{d}u}{\mathrm{d}x}\right]_{i+\frac{1}{2}}-\left[p\frac{\mathrm{d}u}{\mathrm{d}x}\right]_{i-\frac{1}{2}}\right)+\frac{h_{i+1}-h_i}{12}\left[p\frac{\mathrm{d}^3u}{\mathrm{d}x^3}\right]_i+O(h^2)\\&=\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}u}{\mathrm{d}x}\right)\right]_i+\frac{h_{i+1}-h_i}{4}\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(p\frac{\mathrm{d}u}{\mathrm{d}x}\right)\right]_i+\frac{h_{i+1}-h_i}{12}\left[p\frac{\mathrm{d}^3u}{\mathrm{d}x^3}\right]_i+O(h^2).\end{aligned}$$

令 $p_{i-\frac{1}{2}}=p\left(x_{i-\frac{1}{2}}\right), r_i=r(x_i), q_i=q(x_i), f_i=f(x_i)$, 则由 (2.2.3), (2.2.6) 知, 边值问题的解 $u(x)$ 满足方程:

(2.2.7)
$$\begin{aligned}L_hu(x_i)\equiv&-\frac{2}{h_i+h_{i+1}}\left[p_{i+\frac{1}{2}}\frac{u(x_{i+1})-u(x_i)}{h_{i+1}}-p_{i-\frac{1}{2}}\frac{u(x_i)-u(x_{i-1})}{h_i}\right]+\\&\frac{r_i}{h_i+h_{i+1}}\left[u(x_{i+1})-u(x_{i-1})\right]+q_iu(x_i)\\=&f_i+R_i(u),\end{aligned}$$

其中

(2.2.8)
$$R_i(u)=-(h_{i+1}-h_i)\left(\frac{1}{4}\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(p\frac{\mathrm{d}u}{\mathrm{d}x}\right)\right]_i+\frac{1}{12}\left[p\frac{\mathrm{d}^3u}{\mathrm{d}x^3}\right]_i-\frac{1}{2}\left[r\frac{\mathrm{d}^2u}{\mathrm{d}x^2}\right]_i\right)+O(h^2)$$

为差分算子 L_h 的截断误差. 舍去 $R_i(u)$, 便得逼近边值问题 (2.2.1)–(2.2.2) 的差

分方程:

$$(2.2.9) \quad L_h u_i \equiv -\frac{2}{h_i+h_{i+1}}\left[p_{i+\frac{1}{2}}\frac{u(x_{i+1})-u(x_i)}{h_{i+1}}-p_{i-\frac{1}{2}}\frac{u(x_i)-u(x_{i-1})}{h_i}\right]+ \frac{r_i}{h_i+h_{i+1}}(u_{i+1}-u_{i-1})+q_iu_i=f_i,\quad i=1,\cdots,N-1,$$

$$(2.2.10) \quad u_0=\alpha,\quad u_N=\beta.$$

差分方程 (2.2.9) 也可用数值微商公式

$$\left[\frac{\mathrm{d}u}{\mathrm{d}x}\right]_i \approx \frac{u_{i+1}-u_{i-1}}{h_i+h_{i+1}},$$

$$\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}u}{\mathrm{d}x}\right)\right]_i \approx \left(p_{i+\frac{1}{2}}\left[\frac{\mathrm{d}u}{\mathrm{d}x}\right]_{i+\frac{1}{2}}-p_{i-\frac{1}{2}}\left[\frac{\mathrm{d}u}{\mathrm{d}x}\right]_{i-\frac{1}{2}}\right)\Big/\frac{h_i+h_{i+1}}{2}$$
$$\approx \frac{2}{h_i+h_{i+1}}\left(p_{i+\frac{1}{2}}\frac{u_{i+1}-u_i}{h_{i+1}}-p_{i-\frac{1}{2}}\frac{u_i-u_{i-1}}{h_i}\right)$$

代入方程 (2.2.1) 得到. 我们采取前述推导, 是为了导出截断误差公式 (2.2.8).

方程 (2.2.9)–(2.2.10) 是 $N-1$ 阶的线性代数方程组. 若节点次序由左到右排列, 则系数矩阵 $\boldsymbol{A}$ 是三对角矩阵. 由于 r 不恒等于零, 矩阵 $\boldsymbol{A}$ 不对称. 当 $r\equiv 0$ 即 (2.2.1) 对称时, 若网格不均匀, 矩阵 $\boldsymbol{A}$ 也可能不对称. 但可以对称化, 这只要在 (2.2.9) 两端乘以 (h_i+h_{i+1}) 即可. 求解 (2.2.9)–(2.2.10) 就得出解 $u(x)$ 在 x_i 的近似值 u_i.

由方程 (2.2.7), (2.2.9), 截断误差 $R_i(u)$ 可表为

$$(2.2.11) \quad R_i(u)=L_hu(x_i)-L_hu_i=L_h(u(x_i)-u_i).$$

以 $R_h(u)$ 表示由 $R_i(u)$ 定义的网格函数, 则由 (2.2.8) 可知截断误差按 $\|\cdot\|_C$ 或 $\|\cdot\|_0$ 的阶都是 $O(h)$. 当网格均匀, 即 $h_i=h$ $(i=1,2,\cdots,N)$ 时, $\|R_h(u)\|_c$ 或 $\|R_h(u)\|_0$ 的阶提高为 $O(h^2)$. 此时差分方程 (2.2.9) 简化为

$$(2.2.12) \quad L_hu_i=-\frac{1}{h^2}\left[p_{i+\frac{1}{2}}u_{i+1}-(p_{i+\frac{1}{2}}+p_{i-\frac{1}{2}})u_i+p_{i-\frac{1}{2}}u_{i-1}\right]+ r_i\frac{u_{i+1}-u_{i-1}}{2h}+q_iu_i=f_i.$$

这相当于用一阶中心差商、二阶中心差商依次代替方程 (2.2.1) 的一阶微商和二阶微商.

2.2 有限体积法

考虑守恒型微分方程:

$$(2.2.13) \quad Lu=-\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}u}{\mathrm{d}x}\right)+q(x)u=f(x).$$

如果把它看作是分布在一根杆上的稳定温度场方程, 则在 $[a,b]$ 内任一小区间 $[x^{(1)},x^{(2)}]$ 上的热量守恒律具有形式

$$-\int_{x^{(1)}}^{x^{(2)}}\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}u}{\mathrm{d}x}\right)\mathrm{d}x+\int_{x^{(1)}}^{x^{(2)}}qu\mathrm{d}x=\int_{x^{(1)}}^{x^{(2)}}f\mathrm{d}x,$$

或

$$W(x^{(1)})-W(x^{(2)})+\int_{x^{(1)}}^{x^{(2)}}q(x)u\mathrm{d}x=\int_{x^{(1)}}^{x^{(2)}}f\mathrm{d}x, \tag{2.2.14}$$

其中

$$W(x)=p(x)\frac{\mathrm{d}u}{\mathrm{d}x}. \tag{2.2.15}$$

把微分方程 (2.2.13) 写成积分守恒型 (2.2.14) 后, 最高阶微商由二阶降到一阶, 从而可减弱对 p、u 光滑性的要求. 以后会看到, 从积分守恒型方程出发构造差分格式, 便于推广到任意网格和处理第二边值条件.

既然具守恒形式的微分方程反映了物理、力学某些守恒定律, 那么, 我们构造的差分格式也应反映这一基本性质. 现在来构造这种差分格式.

特别于 (2.2.14) 取 $[x^{(1)},x^{(2)}]$ 为对偶单元 $\left[x_{i-\frac{1}{2}},x_{i+\frac{1}{2}}\right]$, 则

$$W\left(x_{i-\frac{1}{2}}\right)-W\left(x_{i+\frac{1}{2}}\right)+\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}qu\mathrm{d}x=\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}f\mathrm{d}x. \tag{2.2.14$'$}$$

考虑到 $p(x)$ 可能有间断点, 此时由 (2.2.15) 进一步差分化是不合适的. 但 “热流量” $W(x)$ 恒连续, 故将 (2.2.15) 改写成

$$\frac{\mathrm{d}u}{\mathrm{d}x}=\frac{W(x)}{p(x)},$$

再沿 $[x_{i-1},x_i]$ 积分, 得

$$u_i-u_{i-1}=\int_{x_{i-1}}^{x_i}\frac{W(x)}{p(x)}\mathrm{d}x,$$

利用中矩形公式, 有

$$W_{i-\frac{1}{2}}\approx a_i\frac{u_i-u_{i-1}}{h_i}, \tag{2.2.16}$$

$$a_i=\left[\frac{1}{h_i}\int_{x_{i-1}}^{x_i}\frac{\mathrm{d}x}{p(x)}\right]^{-1}. \tag{2.2.17}$$

又

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} qu\mathrm{d}x \approx \frac{h_i + h_{i+1}}{2} d_i u_i, \tag{2.2.18}$$

$$d_i = \frac{2}{h_i + h_{i+1}} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q(x)\mathrm{d}x. \tag{2.2.19}$$

将 (2.2.16), (2.2.18) 代到 (2.2.14)′, 即得守恒型差分方程:

$$-\left[a_{i+1}\frac{u_{i+1}-u_i}{h_{i+1}} - a_i\frac{u_i-u_{i-1}}{h_i}\right] + \frac{1}{2}(h_i+h_{i+1})d_iu_i = \frac{1}{2}(h_i+h_{i+1})\phi_i, \tag{2.2.20}$$

$$\phi_i = \frac{2}{h_i + h_{i+1}} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x)\mathrm{d}x. \tag{2.2.21}$$

如果系数 p, q 及右端 f 光滑, 则可用中矩形公式计算 (2.2.17), (2.2.19) 和 (2.2.21), 从而

$$\begin{cases} a_i = p_{i-\frac{1}{2}} = p\left(x_{i-\frac{1}{2}}\right), \\ d_i = q_i = q(x_i), \\ \phi_i = f_i = f(x_i), \end{cases} \tag{2.2.22}$$

也可用梯形公式, 此时

$$\begin{cases} a_i = \dfrac{2p_{i-1}p_i}{p_{i-1}+p_i}, \\ d_i = \dfrac{1}{2}\left(q_{i-\frac{1}{2}} + q_{i+\frac{1}{2}}\right), \\ f_i = \dfrac{1}{2}\left(f_{i-\frac{1}{2}} + f_{i+\frac{1}{2}}\right). \end{cases} \tag{2.2.23}$$

注 2.1 差分方程 (2.2.20) 也适用于具第一类间断系数的微分方程. 此时若系数计算公式 (2.2.22) 或 (2.2.23) 右端的函数在间断点取值, 则应取左右极限的算术平均值. 对于具间断系数的微分方程, 保持守恒形式尤为重要. 例如微分方程

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}u}{\mathrm{d}x}\right) = 0.$$

若把它写成非守恒形式

$$p(x)\frac{\mathrm{d}^2u}{\mathrm{d}x^2} + p'(x)\frac{\mathrm{d}u}{\mathrm{d}x} = 0,$$

再用中心差分格式

$$p_i\frac{u_{i+1}-2u_i+u_{i-1}}{h^2} + \frac{p_{i+1}-p_{i-1}}{2h}\,\frac{u_{i+1}-u_{i-1}}{2h} = 0,$$

则差分解可能不收敛 (参看 [16] 的第 2 章 §2).

2.3 待定系数法

这里我们以一个简单例子说明待定系数法的应用. 考虑边值问题:

$$u''=f,\quad a<x<b, \tag{2.2.24}$$

$$u(a)=u(b)=0. \tag{2.2.25}$$

假定 $\bar{I}_h$ 是均匀网格 ($h_i=h$). 现在构造逼近 (2.2.24)-(2.2.25) 的差分方程. 由 Taylor 展式

$$u(x_{i+1})=u(x_i)+hu'(x_i)+\frac{h^2}{2}u''(x_i)+\frac{h^3}{6}u^{(3)}(x_i)+\frac{h^4}{24}u^{(4)}(x_i)+\frac{h^5}{120}u^{(5)}(x_i)+\frac{h^6}{720}u^{(6)}(\xi_1), \tag{2.2.26}$$

$$u(x_{i-1})=u(x_i)-hu'(x_i)+\frac{h^2}{2}u''(x_i)-\frac{h^3}{6}u^{(3)}(x_i)+\frac{h^4}{24}u^{(4)}(x_i)-\frac{h^5}{120}u^{(5)}(x_i)+\frac{h^6}{720}u^{(6)}(\xi_2), \tag{2.2.27}$$

其中 $\xi_1,\xi_2\in(x_{i-1},x_{i+1})$. 以 α,β,γ 为系数作 $u(x_{i+1}),u(x_i),u(x_{i-1})$ 的一次组合, 得

$$\begin{aligned}&\alpha u(x_{i+1})+\beta u(x_i)+\gamma u(x_{i-1})\\=&(\alpha+\beta+\gamma)u(x_i)+(\alpha-\gamma)hu'(x_i)+\frac{1}{2}(\alpha+\gamma)h^2u''(x_i)+\\&\frac{1}{6}(\alpha-\gamma)h^3u^{(3)}(x_i)+\frac{1}{24}(\alpha+\gamma)h^4u^{(4)}(x_i)+\\&\frac{1}{120}(\alpha-\gamma)h^5u(x_i)+\frac{1}{720}h^6(\alpha u^{(6)}(\xi_1)+\gamma u^{(6)}(\xi_2))\end{aligned}$$

令

$$\alpha+\beta+\gamma=0,$$
$$\alpha-\gamma=0,$$
$$\alpha+\gamma=\frac{2}{h^2}.$$

解之得 $\alpha=\gamma=h^{-2},\beta=-2h^{-2}$. 于是

$$h^{-2}(u(x_{i+1})-2u(x_i)+u(x_{i-1}))=u''(x_i)+\frac{h^2}{12}u^{(4)}(x_i)+O(h^4). \tag{2.2.28}$$

舍去截断误差

$$R_h=\frac{h^2}{12}u^{(4)}(x_i)+O(h^4)$$

便得中心差分格式:

(2.2.29) $$\frac{u_{i+1}-2u_i+u_{i-1}}{h^2}=f_i,\quad i=1,2,\cdots,N-1.$$

如果对方程 (2.2.24) 微商两次, 以 $u^{(4)}=f''$ 代到 (2.2.28) 右端, 还得到精度为 $O(h^4)$ 的格式:

(2.2.30) $$\frac{u_{i+1}-2u_i+u_{i-1}}{h^2}=f_i+\frac{h^2}{12}f_i''.$$

2.4 边值条件的处理

最简单的第一边值条件已处理过了, 现在处理第二、第三边值条件:

$(2.2.31)_1$ $$u'(a)=\alpha_0u(a)+\alpha_1,$$

$(2.2.31)_2$ $$u'(b)=\beta_0u(b)+\beta_1.$$

最容易想到的是用数值微商公式

$$u'(a)\approx\frac{u_1-u_0}{h_1},\quad u'(b)\approx\frac{u_N-u_{N-1}}{h_N}$$

代替 $(2.2.31)_1$、$(2.2.31)_2$ 中的微商. 但这样处理有两个缺点: 一是截断误差的阶比内点低, 例如对均匀网格, 内点的截断误差为 $O(h^2)$, 界点的截断误差的阶为 $O(h)$; 二是可能会破坏差分方程 (2.2.20) 的对称性. 为此我们用有限体积法, 像推导内点差分方程那样导出近似边值条件.

因为 $p(x)>0$, 不失一般性可将边值条件 $(2.2.31)_1$、$(2.2.31)_2$ 写成形式

$(2.2.32)_1$ $$-p(a)u'(a)=\alpha_0u(a)+\alpha_1,$$

$(2.2.32)_2$ $$-p(b)u'(b)=\beta_0u(b)+\beta_1.$$

于积分守恒形式 (2.2.14) 中取 $x^{(1)}=x_0=a$, $x^{(2)}=x_{\frac{1}{2}}$, 得

$$W(a)-W\left(x_{\frac{1}{2}}\right)+\int_{x_0}^{x_{\frac{1}{2}}}qu\mathrm{d}x=\int_{x_0}^{x_{\frac{1}{2}}}f\mathrm{d}x.$$

而

$$W(a)=p(x)\left.\frac{\mathrm{d}u}{\mathrm{d}x}\right|_{x=a}=-(\alpha_0u_0+\alpha_1),$$

故

(2.2.33) $$-W(x_{\frac{1}{2}})+\int_{x_0}^{x_{\frac{1}{2}}}qu\mathrm{d}x=(\alpha_0u_0+\alpha_1)+\int_{x_0}^{x_{\frac{1}{2}}}f\mathrm{d}x.$$

由 (2.2.16) 得

$$W(x_{\frac{1}{2}}) \approx a_1 \frac{u_1 - u_0}{h_1}, a_1 = \left[\frac{1}{h_1}\int_{x_0}^{x_1} \frac{\mathrm{d}x}{p(x)}\right]^{-1}. \tag{2.2.34}$$

又

$$\int_{x_0}^{x_{\frac{1}{2}}} qu\mathrm{d}x \approx \frac{h_1}{2} d_0 u_0, d_0 = \frac{2}{h_1}\int_{x_0}^{x_{\frac{1}{2}}} q\mathrm{d}x, \tag{2.2.35}$$

$$\int_{x_0}^{x_{\frac{1}{2}}} f\mathrm{d}x = \frac{h_1}{2}\phi_0, \phi_0 = \frac{2}{h_1}\int_{x_0}^{x_{\frac{1}{2}}} f\mathrm{d}x. \tag{2.2.36}$$

以 (2.2.34)~(2.2.36) 代到 (2.2.33), 得

$$-a_1\frac{u_1 - u_0}{h_1} + \left(-\alpha_0 + \frac{h_1}{2}d_0\right)u_0 - \left(\alpha_1 + \frac{h_1}{2}\phi_0\right) = 0. \tag{2.2.37}$$

类似地可导出逼近 $(2.2.32)_2$ 的差分方程.

可以证明, 当网格均匀且系数光滑时, 差分方程 (2.2.20) 逼近 (2.2.13) 的阶为 $O(h^2)$, 边界差分方程 (2.2.37) 逼近 $(2.2.31)_1$ 的阶亦为 $O(h^2)$.

习　　题

1. 用有限体积法导出逼近微分方程 (2.2.1) 的差分方程.
2. 构造逼近

$$(pu'')'' + (qu')' + ru = f, \quad 于\ (a, b),$$
$$u(a) = u'(a) = 0, \quad u(b) = u'(b) = 0$$

的中心差分格式.

§3 矩形网的差分格式

考虑 Poisson 方程:

$$-\Delta u = f(x, y), \quad (x, y) \in G, \tag{2.3.1}$$

G 是 xy 平面上一有界区域, 其边界 Γ 为分段光滑曲线. 在 Γ 上 u 满足下列边值条件之一:

$$u\big|_{\Gamma} = \alpha(x, y) \quad (第一边值条件), \tag{2.3.1$_1$}$$

$$\left.\frac{\partial u}{\partial n}\right|_{\Gamma} = \beta(x, y) \quad (第二边值条件), \tag{2.3.1$_2$}$$

$$\frac{\partial u}{\partial n} + ku\big|_{\Gamma} = \gamma(x, y) \quad (第三边值条件), \tag{2.3.1$_3$}$$

其中 $f(x,y),\alpha(x,y),\beta(x,y),\gamma(x,y)$ 及 $k(x,y)$ 都是连续函数, $k \geqslant 0$. 本节讨论逼近方程 (2.3.1) 及相应边值条件的差分格式. 原则上, 前节的方法都可推广到二维边值问题, 但会遇到新的困难. 例如随着维数的增加, 求解域的几何形状会更复杂, 如何作网格剖分及处理好边值条件, 就是一个重要而困难的问题.

3.1 五点差分格式

取定沿 x 轴和 y 轴的步长 h_1 和 h_2, $h=(h_1^2+h_2^2)^{\frac{1}{2}}$. 作两族与坐标轴平行的直线:

$$x=ih_1, \quad i=0,\pm 1,\cdots,$$
$$y=jh_2, \quad j=0,\pm 1,\cdots.$$

两族直线的交点 (ih_1,jh_2) 称为**网点或节点**, 记为 (x_i,y_j) 或 (i,j). 两个节点 (x_i,y_j) 和 $(x_{i'},y_{j'})$ 说是相邻, 如果

$$\left|\frac{x_i-x_{i'}}{h_1}\right|+\left|\frac{y_j-y_{j'}}{h_2}\right|=1 \text{ 或 } |i-i'|+|j-j'|=1.$$

以 $G_h=\{(x_i,y_i)\in G\}$ 表示所有属于 G 内部的节点集合, 并称如此的节点为**内点**. 以 $\varGamma_h$ 表示网线 $x=x_i$ 或 $y=y_j$ 与 $\varGamma$ 的交点集合, 并称如此的点为**界点**. 令 $\overline{G}_h=G_h\bigcup\varGamma_h$, 则 $\overline{G}_h$ 就是代替域 $\overline{G}=G\bigcup\varGamma$ 的网点集合. 若内点 (x_i,y_j) 的四个相邻点都属于 G_h, 就称为**正则内点**; 否则称为**非正则内点**. 图 3.1 中打 "○" 号的点为正则内点, 打 "×" 号的点为非正则内点, 打 "•" 号的点为界点.

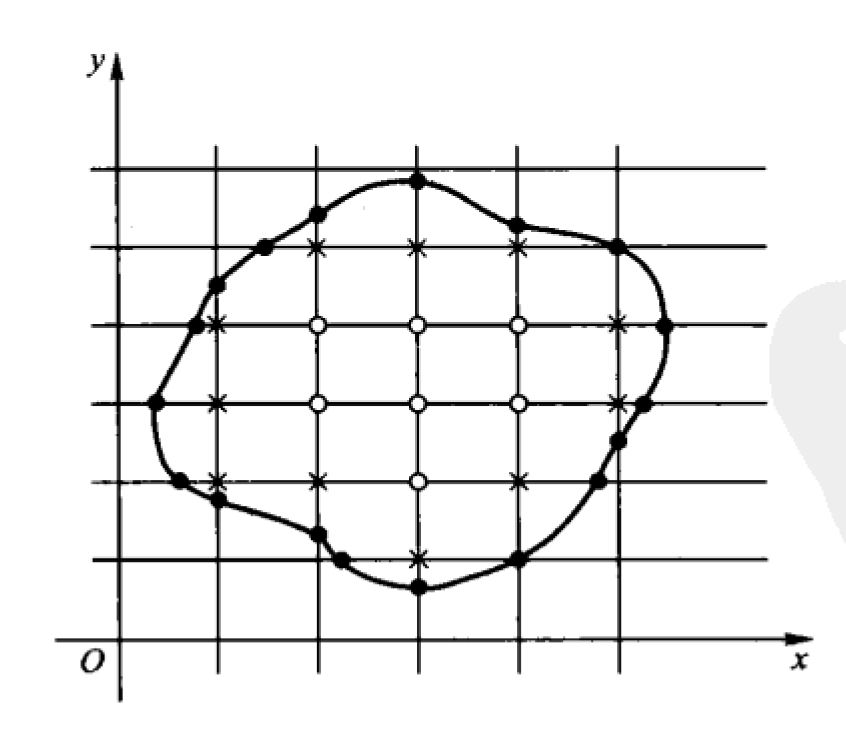


图 3.1

现在假定 (x_i,y_j) 为正则内点. 沿 x, y 方向分别用二阶中心差商代替 u_{xx},

u_{yy}, 则得

$$-\Delta_h u_{ij} = -\left[\frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h_1^2} + \frac{u_{i,j+1} - 2u_{ij} + u_{i,j-1}}{h_2^2}\right] = f_{ij}, \tag{2.3.2}$$

式中 u_{ij} 表示节点 (i,j) 上的网函数值. 若以 u_h, f_h 表示网格函数, $u_h(x_i,y_j) = u_{ij}$, $f_h(x_i,y_j) = f_{ij} = f(x_i,y_j)$, 则差分方程 (2.3.2) 可简写成

$$-\Delta_h u_h = f_h. \tag{2.3.2$'$}$$

利用 Taylor 展式,

$$\begin{aligned}(2.3.3)\qquad &\frac{u(x_{i+1},y_j) - 2u(x_i,y_j) + u(x_{i-1},y_j)}{h_1^2}\\ &= \frac{\partial^2 u(x_i,y_j)}{\partial x^2} + \frac{h_1^2}{12}\frac{\partial^4 u(x_i,y_j)}{\partial x^4} + \frac{h_1^4}{360}\frac{\partial^6 u(x_i,y_j)}{\partial x^6} + O(h_1^6),\\ (2.3.4)\qquad &\frac{u(x_i,y_{j+1}) - 2u(x_i,y_j) + u(x_i,y_{j-1})}{h_2^2}\\ &= \frac{\partial^2 u(x_i,y_j)}{\partial y^2} + \frac{h_2^2}{12}\frac{\partial^4 u(x_i,y_j)}{\partial y^4} + \frac{h_2^4}{360}\frac{\partial^6 u(x_i,y_j)}{\partial y^6} + O(h_2^6),\end{aligned}$$

可得差分算子 $-\Delta_h$ 的截断误差

$$\begin{aligned}R_{ij}(u) &= \Delta u(x_i,y_j) - \Delta_h u(x_i,y_j)\\ &= -\frac{1}{12}\left[h_1^2\frac{\partial^4 u}{\partial x^4} + h_2^2\frac{\partial^4 u}{\partial y^4}\right]_{ij} + O(h^4)\\ &= O(h^2),\end{aligned} \tag{2.3.5}$$

其中 u 是方程 (2.3.1) 的光滑解.

由于差分方程 (2.3.2) 中只出现 u 在 (i,j) 及其四个邻点上的值, 故称为**五点差分格式**, 其图式如图 3.2.

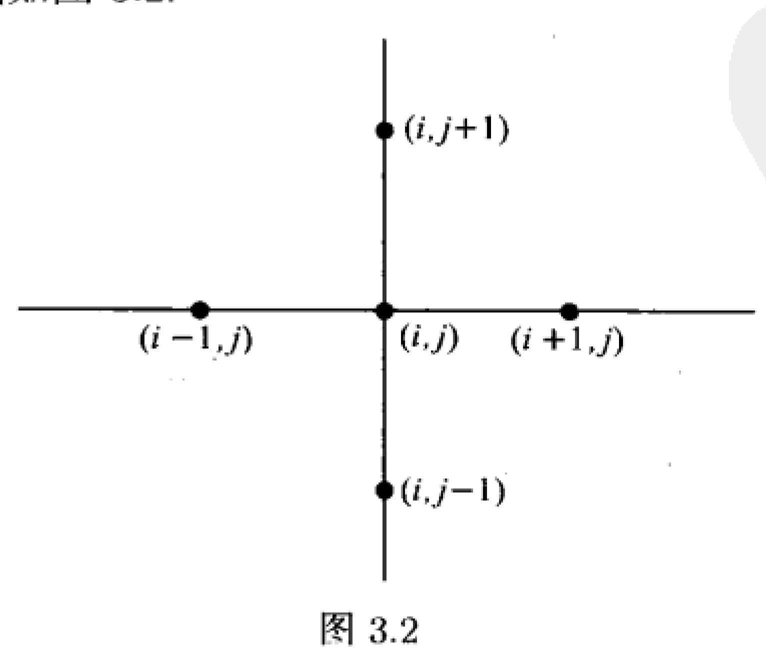


图 3.2

特别取正方形网格: $h_1 = h_2 = h$, 则差分方程 (2.3.2) 简化为

$$u_{ij} - \frac{1}{4}(u_{i-1,j} + u_{i,j-1} + u_{i+1,j} + u_{i,j+1}) = \frac{h^2}{4} f_{ij}. \tag{2.3.6}$$

若 $f \equiv 0$ (Laplace 方程), 则

$$u_{ij} = \frac{1}{4}(u_{i-1,j} + u_{i,j-1} + u_{i+1,j} + u_{i,j+1}). \tag{2.3.7}$$

注 3.1 若将 (2.3.3), (2.3.4) 两式相加, 则得

$$\begin{aligned}
\Delta_h u(x_i, y_j) &= \Delta u(x_i, y_j) + \frac{1}{12}\left(h_1^2 \frac{\partial^4 u(x_i, y_j)}{\partial x^4} + h_2^2 \frac{\partial^4 u(x_i, y_j)}{\partial y^4}\right) + O(h^4) \\
&= \Delta u(x_i, y_j) + \frac{1}{12}\left(h_1^2 \frac{\partial^2}{\partial x^2} + h_2^2 \frac{\partial^2}{\partial y^2}\right)\left(\frac{\partial^2 u(x_i, y_j)}{\partial x^2} + \frac{\partial^2 u(x_i, y_j)}{\partial y^2}\right) - \\
&\quad \frac{h_1^2 + h_2^2}{12} \frac{\partial^4 u(x_i, y_j)}{\partial x^2 \partial y^2} + O(h^4) \\
&= -f(x_i, y_j) - \frac{1}{12}\left(h_1^2 \frac{\partial^2 f(x_i, y_j)}{\partial x^2} + h_2^2 \frac{\partial^2 f(x_i, y_j)}{\partial y^2}\right) - \\
&\quad \frac{h_1^2 + h_2^2}{12} \frac{\partial^4 u(x_i, y_j)}{\partial x^2 \partial y^2} + O(h^4).
\end{aligned}$$

又

$$\begin{aligned}
\frac{\partial^4 u(x_i, y_j)}{\partial x^2 \partial y^2} &= \frac{u''_{xx}(x_i, y_{j+1}) - 2u''_{xx}(x_i, y_j) + u''_{xx}(x_i, y_{j-1})}{h_2^2} + O(h_2^2) \\
&= \frac{1}{h_1^2 h_2^2}[u(x_{i+1}, y_{j+1}) - 2u(x_i, y_{j+1}) + \\
&\quad u(x_{i-1}, y_{j+1}) - 2(u(x_{i+1}, y_j) - 2u(x_i, y_j) + \\
&\quad u(x_{i-1}, y_j)) + u(x_{i+1}, y_{j-1}) - \\
&\quad 2u(x_i, y_{j-1}) + u(x_{i-1}, y_{j-1})] + O(h^2).
\end{aligned}$$

因此

$$\begin{aligned}
&\Delta_h u(x_i, y_j) + \frac{1}{12}[4u(x_i, y_j) - 2(u(x_{i-1}, y_j) + u(x_i, y_{j-1}) + \\
&u(x_{i+1}, y_j) + u(x_i, y_{j+1})) + u(x_{i-1}, y_{j-1}) + u(x_{i+1}, y_{j-1}) + \\
&u(x_{i+1}, y_{j+1}) + u(x_{i-1}, y_{j+1})](h_1^2 + h_2^2)/h_1^2 h_2^2 \\
= &-f(x_i, y_j) - \frac{1}{12}\left(h_1^2 \frac{\partial^2 f(x_i, y_j)}{\partial x^2} + h_2^2 \frac{\partial^2 f(x_i, y_j)}{\partial y^2}\right) + O(h^4).
\end{aligned}$$

舍去截断误差项, 便得到逼近 Poisson 方程的**九点差分格式**:

$$\begin{aligned}
&-\Delta_h u_{ij} - \frac{1}{12}[4u_{ij} - 2(u_{i-1,j} + u_{i,j-1} + u_{i+1,j} + u_{i,j+1}) + \\
&u_{i-1,j-1} + u_{i+1,j-1} + u_{i+1,j+1} + u_{i-1,j+1}](h_1^2 + h_2^2)/h_1^2 h_2^2 \\
= &f_{ij} + \frac{1}{12}[h_1^2 f''_{xx}(x_i, y_j) + h_2^2 f''_{yy}(x_i, y_j)],
\end{aligned}$$

其截断误差的阶为 $O(h^4)$.

现在用有限体积法推导五点格式. 为此我们需要作对偶部分. 记 $x_{i-\frac{1}{2}} = \left(i-\frac{1}{2}\right)h_1, y_{j-\frac{1}{2}} = \left(j-\frac{1}{2}\right)h_2$. 作两族平行于坐标轴的直线 $x = x_{i-\frac{1}{2}}$ 和 $y = y_{j-\frac{1}{2}}$, $i,j = 0,\pm1,\cdots$, 其交点属于 G 内部者为对偶部分的内点, 直线与边界 Γ 的交点为对偶剖分的界点. 对于任一正则内点 (x_i,y_j), 考虑对偶剖分的网点: $A(x_{i-\frac{1}{2}},y_{j-\frac{1}{2}})$, $B(x_{i+\frac{1}{2}},y_{j-\frac{1}{2}})$, $C(x_{i+\frac{1}{2}},y_{j+\frac{1}{2}})$, $D(x_{i-\frac{1}{2}},y_{j+\frac{1}{2}})$, 用 G_{ij} 表示以 A,B,C,D 为顶点的矩形域, 称为控制体积, ∂G_{ij} 为其边界 (参看图 3.3). 于 G_{ij}

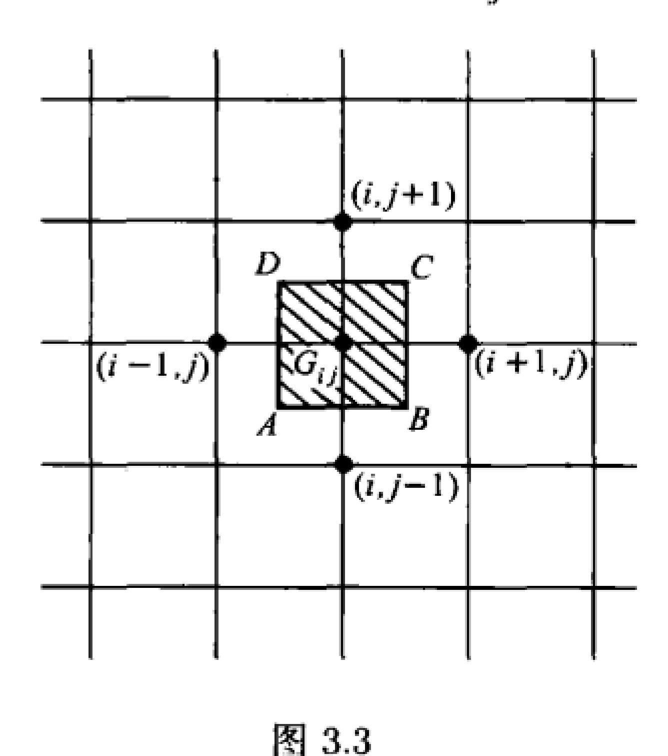


图 3.3

积分 Poisson 方程 (2.3.1), 并利用 Green 公式, 得到 Poisson 方程的积分守恒形式:

$$-\int_{\partial G_{ij}} \frac{\partial u}{\partial n}\mathrm{d}s = \iint_{G_{ij}} f\mathrm{d}x\mathrm{d}y. \tag{2.3.8}$$

式中 $\dfrac{\partial u}{\partial n}$ 表示 u 沿矩形 ∂G_{ij} 的外法向导数. 用中矩形公式代替沿四边的线积分, 再用中心差商代替外法向导数, 则

$$\int_{\partial G_{ij}} \frac{\partial u}{\partial n}\mathrm{d}s \approx \frac{u_{i,j-1}-u_{ij}}{h_2}h_1 + \frac{u_{i+1,j}-u_{ij}}{h_1}h_2 + \frac{u_{i,j+1}-u_{ij}}{h_2}h_1 + \frac{u_{i-1}-u_{ij}}{h_1}h_2.$$

以之代到 (2.3.8), 并除以 h_1h_2, 即得五点差分格式:

$$-\left[\frac{u_{i+1,j}-2u_{ij}+u_{i-1,j}}{h_1^2} + \frac{u_{i,j+1}-2u_{ij}+u_{i,j-1}}{h_2^2}\right] = \phi_{ij}, \tag{2.3.9}$$

$$\phi_{ij} = \frac{1}{h_1h_2}\iint_{G_{ij}} f\mathrm{d}x\mathrm{d}y \approx f_{ij},$$

它和 (2.3.2) 一致.

3.2 边值条件的处理

先讨论第一边值条件

$$u\big|_{\Gamma} = \alpha(x, y). \tag{2.3.10}$$

以 G_h^* 表示非正则内点集合, Γ_h 表示界点集合. 因 $\Gamma_h \subset \Gamma$, 当 $(\overline{x}_i, \overline{y}_j) \in \Gamma_h$ 时便令

$$u_{ij} = \alpha(\overline{x}_i, \overline{y}_j). \tag{2.3.11}$$

当 $(\overline{x}_i, \overline{y}_j) \in G_h^*$ 即它为非正则内点时, 将它和正则内点一样, 在 $(\overline{x}_i, \overline{y}_j)$ 建立逼近 Poisson 方程的不等距差分格式. 例如在节点 0 (参看图 3.4) 有

$$-\left[\frac{1}{\overline{h}_1}\left(\frac{u_1 - u_0}{h_1} - \frac{u_0 - u_3}{h_1^-}\right) + \frac{1}{h_2^2}(u_2 - 2u_0 + u_4)\right] = f_0, \tag{2.3.12}$$

$$\overline{h}_1 = \frac{1}{2}(h_1 + h_1^-), h_1^- = x_0 - x_3.$$

其截断误差的阶为 $O(h)$.

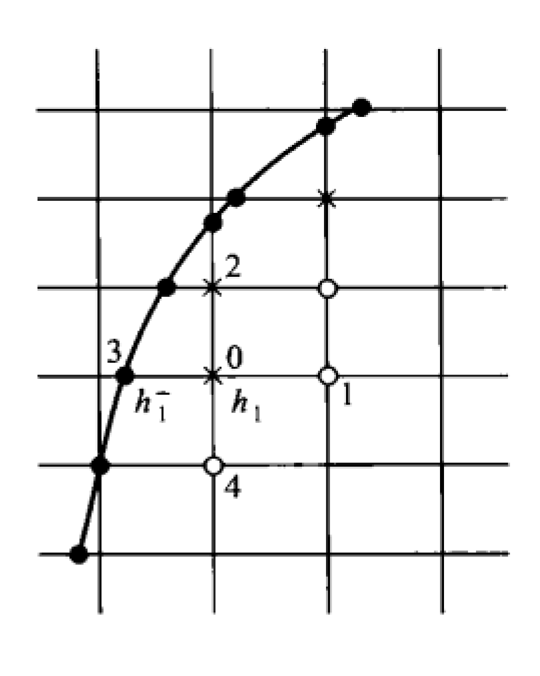


图 3.4

按 (2.3.12) 处理边值条件有一个缺点, 即破坏了对称正定性, 而这一性质是五点差分格式所固有的. 为了保持对称正定性, 可用

$$-\left[\frac{1}{h_1}\left(\frac{u_1 - u_0}{h_1} - \frac{u_0 - u_3}{h_1^-}\right) + \frac{1}{h_2^2}(u_2 - 2u_0 + u_4)\right] = f_0 \tag{2.3.13}$$

代替 (2.3.12). 此时截断误差的阶降为 $O(1)$. 尽管如此, 仍可证明, 差分解的收敛阶仍是 $O(h^2)$ (按最大模) (见 [16] 的第三章 §4).

今讨论第二、第三边值条件:

$$\frac{\partial u}{\partial n} + ku\big|_{\Gamma} = \gamma. \tag{2.3.14}$$

我们用有限体积法构造 (2.3.14) 的差分逼近. 我们只讨论一种特殊情况, 假定 Γ_h 中的节点 (界点) 是两族网线的交点 (作网格时可要求界点为两族网格交点). 如

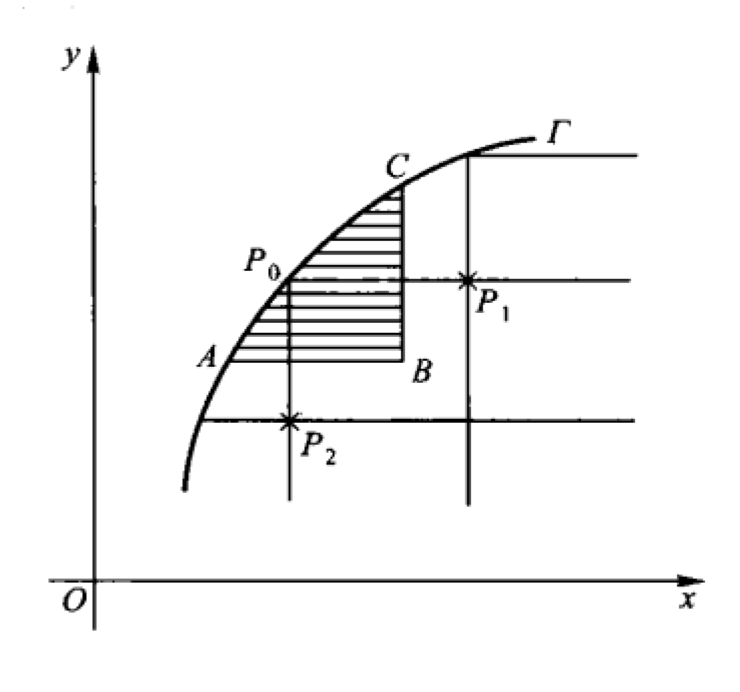


图 3.5

图 3.5, $P_0(x_{i_0}, y_{j_0})$ 是界点, $P_1(x_{i_0+1}, y_{j_0})$ 和 $P_2(x_{i_0}, y_{j_0-1})$ 是与之相邻的内点. 过 $(x_{i_0+\frac{1}{2}}, y_{j_0})$, $(x_{i_0}, y_{j_0-\frac{1}{2}})$ 分别作与 y 轴和 x 轴平行的直线, 它们与外边界 Γ 一起截出一曲边三角形 $\widetilde{\triangle}ABC$. 于 $\widetilde{\triangle}ABC$ 积分 (2.3.1) 两端, 并利用 Green 公式, 得

$$-\int_{\widetilde{\triangle}ABCA} \frac{\partial u}{\partial n} \mathrm{d}s = \iint\limits_{\widetilde{\triangle}ABC} f \mathrm{d}x\mathrm{d}y. \tag{2.3.15}$$

而

$$\begin{aligned}
\int_{\overline{AB}} \frac{\partial u}{\partial n} \mathrm{d}s &\approx \frac{u_{P_2} - u_{P_0}}{h_2} \cdot \overline{AB}, \\
\int_{\overline{BC}} \frac{\partial u}{\partial n} \mathrm{d}s &\approx \frac{u_{P_1} - u_{P_0}}{h_1} \cdot \overline{BC}, \\
\int_{\widehat{CA}} \frac{\partial u}{\partial n} \mathrm{d}s &= \int_{\widehat{CA}} (\gamma - ku) \mathrm{d}s \\
&\approx (\gamma_{P_0} - k_{P_0} u_{P_0}) \times \widehat{CA},
\end{aligned}$$

以之代到 (2.3.15) 即得逼近 (2.3.14) 的差分方程:

$$-\left[\frac{u_{p_2} - u_{p_0}}{h_2} \cdot \overline{AB} + \frac{u_{p_1} - u_{p_0}}{h_1} \cdot \overline{BC} + (\gamma_{p_0} - k_{p_0} u_{p_0}) \times \widehat{CA}\right] = \iint\limits_{\widetilde{\triangle}ABC} f \mathrm{d}x\mathrm{d}y. \tag{2.3.16}$$

可见用有限体积法处理第二、第三边值条件特别方便. 对非正则内点, 仍可像前面一样建立一形如 (2.3.12) 或 (2.3.13) 的方程.

3.3　极坐标形式的差分格式

如果求解域是圆域、环形域或扇形域, 则采用极坐标是方便的. 此时 Poisson 方程形如

$$-\Delta_{r,\theta}u = -\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2}\right] = f(r,\theta), \tag{2.3.17}$$

其中 $r=\sqrt{x^2+y^2}$, $\text{tg}\theta = y/x$. 整个 xy 平面变成 $r\theta$ 平面上的半带形域 $\{0\leqslant r<\infty, 0\leqslant\theta\leqslant 2\pi\}$.

方程 (2.3.17) 的系数于 $r=0$ 奇异, 因此只当 $r>0$ 时有意义. 在 $r=0$ 需补充 u 为光滑的条件 (原点是可去奇点). 由 $u(0,\theta)=u(r,\theta)-ru_r(r,\theta)+O(r^2)$, $r\to 0$, 则知

$$\lim_{r\to 0^+} r\frac{\partial u}{\partial r} = 0. \tag{2.3.18}$$

对变量 r, θ 分别取等步长 h_r 和 h_θ. 令

$$\begin{aligned} &r_i = (i+0.5)h_r, i=0,1,2,\cdots,\\ &\theta_j = (j+1)h_\theta, j=0,1,\cdots,J-1, h_\theta = 2\pi/J, \end{aligned}$$

则半带形域的网格节点为 (r_i,θ_j), 它们在 $r\theta$ 平面上的分布如图 3.6.

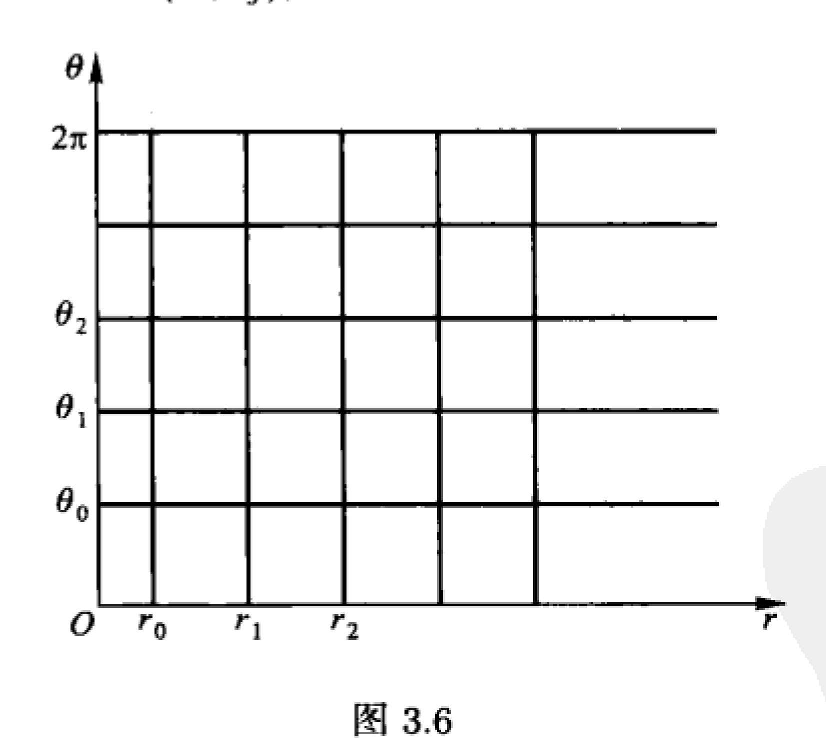


图 3.6

对任一点 $(r_i,\theta_j)(i\geqslant 1)$, 用中心差商公式

$$\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right)\right]_{(r_i,\theta_j)} \approx \frac{1}{r_i}\frac{r_{i+\frac{1}{2}}u_{i+1,j}-(r_{i+\frac{1}{2}}+r_{i-\frac{1}{2}})u_{ij}+r_{i-\frac{1}{2}}u_{i-1,j}}{h_r^2},$$

$$\left[\frac{1}{r^2}\frac{\partial^2 u}{\partial\theta^2}\right]_{(r_i,\theta_j)} \approx \frac{1}{r_i^2}\frac{u_{i,j+1}-2u_{ij}+u_{i,j-1}}{h_\theta^2}(u_{i0}=u_{iJ}, u_{i,-1}=u_{i,J-1})$$

代到 (2.3.17), 则就 $i \geqslant 1$ 得到逼近它的差分方程:

(2.3.19)
$$-\left[\frac{1}{r_i}\frac{r_{i+\frac{1}{2}}u_{i+1,j}-(r_{i+\frac{1}{2}}+r_{i-\frac{1}{2}})u_{ij}+r_{i-\frac{1}{2}}u_{i-1,j}}{h_r^2}+\frac{1}{r_i^2}\frac{u_{i,j+1}-2u_{ij}+u_{i,j-1}}{h_\theta^2}\right]$$
$$=f(r_i,\theta_j).$$

今用有限体积法导出点 (r_0,θ_j) 的差分方程. 以 r 乘方程 (2.3.17) 两端, 并对 r 由 ε 到 h_r 积分, 对 θ 由 $\theta_{j-\frac{1}{2}}$ 到 $\theta_{j+\frac{1}{2}}$ 积分, 然后令 $\varepsilon\to 0$, 则利用条件 (2.3.18), 可得

$$-h_r\int_{\theta_{j-\frac{1}{2}}}^{\theta_{j+\frac{1}{2}}}\frac{\partial}{\partial r}u(h_r,\theta)\mathrm{d}\theta-\int_0^{h_r}\frac{1}{r}\left(\frac{\partial}{\partial\theta}u\left(r,\theta_{j+\frac{1}{2}}\right)-\frac{\partial}{\partial\theta}u(r,\theta_{j-\frac{1}{2}})\right)\mathrm{d}r$$
$$=\int_0^{h_r}r\mathrm{d}r\int_{\theta_{j-\frac{1}{2}}}^{\theta_{j+\frac{1}{2}}}f(r,\theta)\mathrm{d}\theta.$$

用中矩形公式代替上述积分, 则

$$-h_rh_\theta\frac{\partial}{\partial r}(h_r,\theta_j)-2\left(\frac{\partial}{\partial\theta}u\left(\frac{h_r}{2},\theta_{j+\frac{1}{2}}\right)-\frac{\partial}{\partial\theta}u\left(\frac{h_r}{2},\theta_{j-\frac{1}{2}}\right)\right)\approx\frac{1}{2}h_r^2h_\theta f\left(\frac{h_r}{2},\theta_j\right).$$

再用中心差商代替相应微商, 就得到点 (r_0,θ_j) 的差分方程:

$$-h_rh_\theta\frac{u_{1j}-u_{0j}}{h_r}-2\frac{u_{0,j+1}-2u_{0j}+u_{0,j-1}}{h_\theta}=\frac{1}{2}h_r^2h_\theta f_{0j}$$

或 (两端除以 $\frac{1}{2}h_r^2h_\theta$)

(2.3.20)
$$-\frac{2}{h_r}\frac{u_{1j}-u_{0j}}{h_r}-\frac{4}{h_r^2}\frac{u_{0,j+1}-2u_{0j}+u_{0,j-1}}{h_\theta^2}=f_{0j}.$$

这就给出了逼近 (2.3.17) 的差分方程 (2.3.19), (2.3.20).

习　　题

1. 用有限体积法构造逼近方程

(2.3.21)
$$-\nabla(k\nabla u)=-\left[\frac{\partial}{\partial x}\left(k\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial y}\left(k\frac{\partial u}{\partial y}\right)\right]=f$$

的第一边值问题的五点差分格式, 这里 $k=k(x,y)\geqslant k_{\min}>0$.

2. 用有限体积法构造逼近方程 (2.3.21) 的第二边值问题的五点差分格式.

3. 设步长 $h_1=h_2=h$. 记

$$\Diamond u_{ij}=u_{i+1,j}+u_{i,j+1}+u_{i-1,j}+u_{i,j-1}-4u_{ij},$$
$$\Box u_{ij}=u_{i+1,j+1}+u_{i-1,j+1}+u_{i-1,j-1}+u_{i+1,j-1}-4u_{ij}.$$

证明逼近 Laplace 方程 $\Delta u = 0$ 的差分方程

$$\frac{1}{6h^2}[4\Diamond u_{i,j} + \Box u_{i,j}] = 0$$

的截断误差的绝对值

$$|R_{i,j}(u)| = \frac{40h^6}{3 \cdot 8!}\theta M_8,$$

其中 M_8 是 u 的八阶偏导数的绝对值于考虑区域的上确界, $|\theta| \leqslant 1$.

§4 三角网的差分格式

利用有限体积法, 可将矩形网的差分格式推广到三角网, 得到三角网的差分格式, 文献上也称之为广义差分法 (见 [9]). 三角网的差分格式具有网格灵活, 边值条件容易处理等优点, 特别它还保持积分守恒 (质量守恒), 所以受到应用部门的欢迎.

考虑有界域 G 上的 Poisson 方程

$$-\Delta u = f, \tag{2.4.1}$$

在边界 Γ 上满足第一、第二或第三边值条件. 我们先对 G 做三角剖分. 如图 4.1, 在 Γ 上取一系列节点, 以它们为顶点做逼近 Γ 的闭折线 $\widetilde{\Gamma}$, 设 $\widetilde{G}$ 为由 $\widetilde{\Gamma}$ 围成的逼近 G 的多边形域, 然后把 $\widetilde{G}$ 分割成有限个小三角形之和, 使不同三角形无重叠的内部区域, 任一三角形的顶点不属于其他三角形边的内部. 此外, 我们还要求这些三角形的内角不大于 90°. 这样, 就把 $\widetilde{G}$ 分割成一三角网, 称为 G 的**三角剖分**. 组成剖分的小三角形也称为**三角单元**.

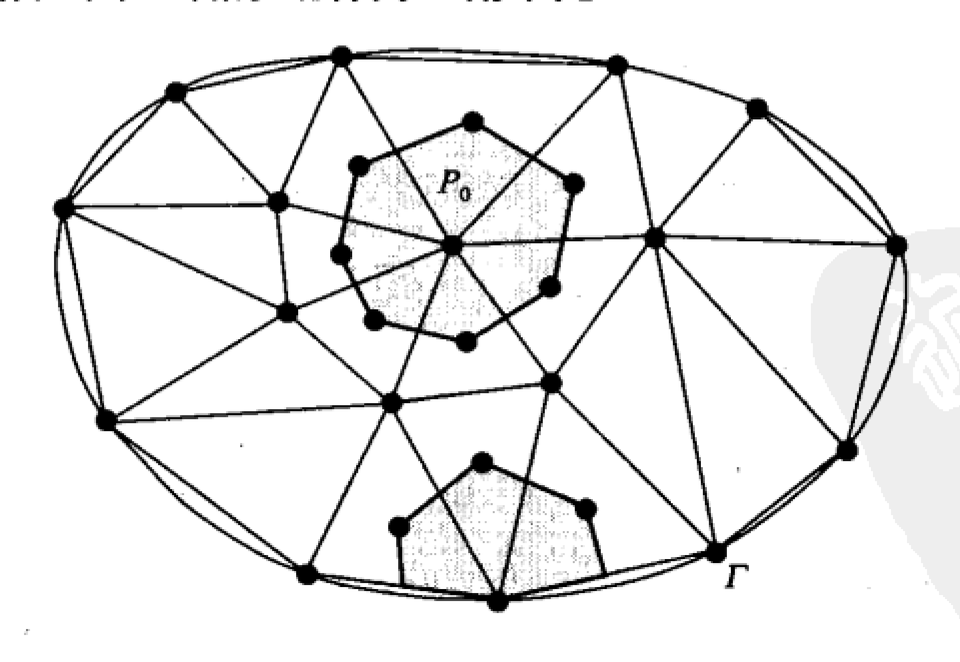


图 4.1

对于任一节点, 考虑所有以它为顶点的三角单元和以它为端点的三角形的边. 过每边作中垂线, 它们依次交于相应三角形的外心, 从而得到围绕该节点的

小多边形域, 称为**对偶单元** (也叫控制体积). 这些对偶单元全体构成区域 G 的一个新的网格剖分, 称为**对偶剖分**. 如图 4.2 (a), 内点 $P_0 \in G$ 的对偶单元是六边形域, P_0 在六边形内部. 若 P_0 是界点, 即 $P_0 \in \Gamma$, 则 P_0 是其对偶单元的一个顶点. (参看图 4.2(b))

现就每一内点建立差分方程. 如图 4.2(a), 设 P_0 是内点, $P_1, P_2, \cdots, P_6$ 是和 P_0 相邻的节点, q_i 为三角形 $P_0P_iP_{i+1}$ ($P_7 = P_1$, 下同) 的外心, m_i 是线段 $\overline{P_0P_i}$ 的中点, G_0 是由六边形 $q_1q_2\cdots q_6$ 围成的对偶单元. 在子域 G_0 积分 (2.4.1), 得

$$-\iint\limits_{G_0} \Delta u \mathrm{d}x\mathrm{d}y = \iint\limits_{G_0} f\mathrm{d}x\mathrm{d}y.$$

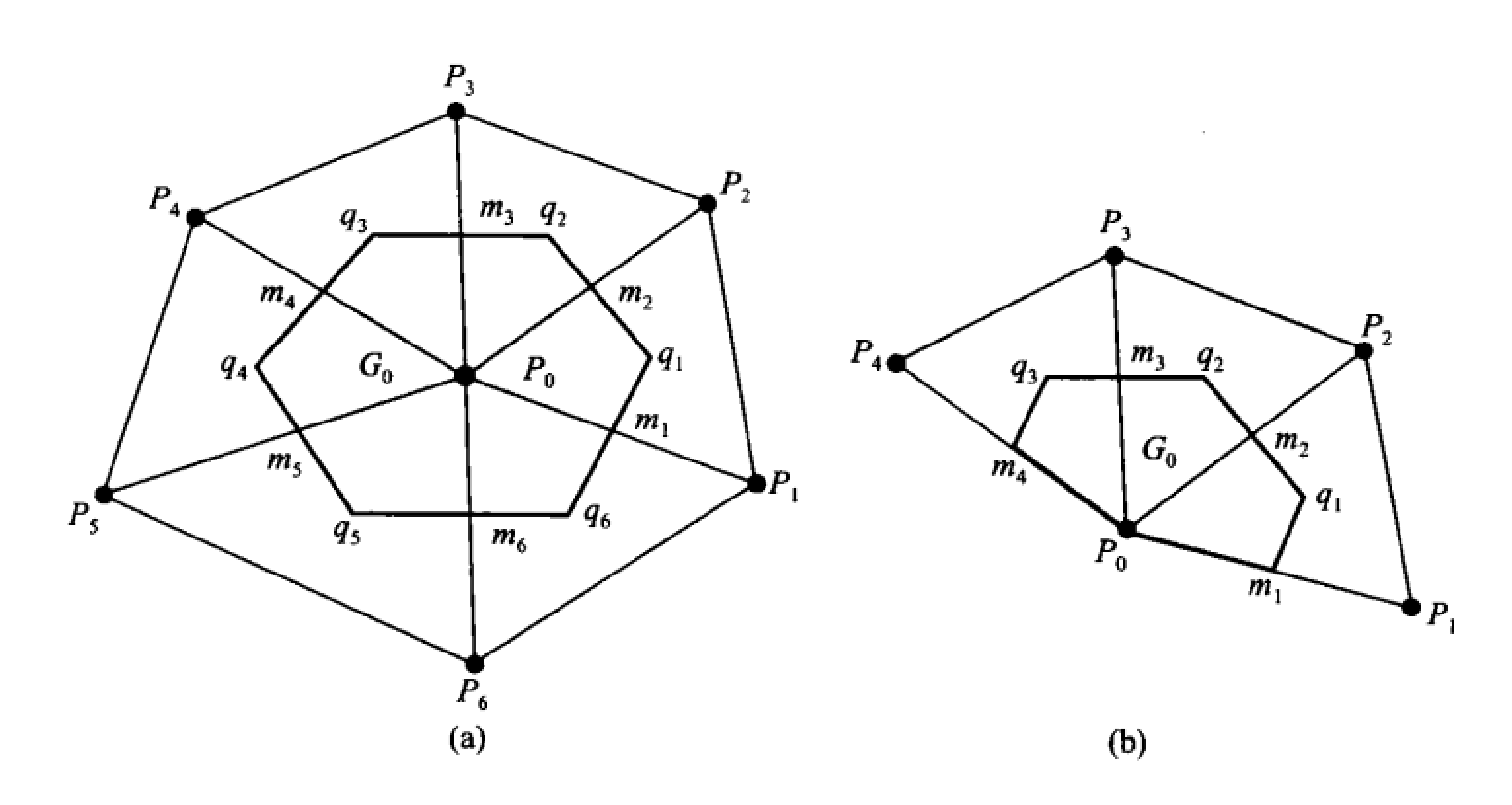


图 4.2

利用 Green 公式, 可将上式改写成

$$-\int_{\partial G_0} \frac{\partial u}{\partial \boldsymbol{n}} \mathrm{d}s = \iint\limits_{G_0} f\mathrm{d}x\mathrm{d}y. \tag{2.4.2}$$

∂G_0 是 G_0 的边界, $\boldsymbol{n}$ 是 ∂G_0 的单位外法向. 注意

$$\int_{\partial G_0} \frac{\partial u}{\partial \boldsymbol{n}} \mathrm{d}s = \sum_{i=1}^{6} \int_{\overline{q_iq_{i+1}}} \frac{\partial u}{\partial \boldsymbol{n}} \mathrm{d}s = \sum_{i=1}^{6} \frac{\overline{q_iq_{i+1}}}{\overline{P_0P_{i+1}}} \cdot [u(P_{i+1}) - u(P_0)] + m(G_0)R_{G_0}(u), \ \ (q_7 = q_1) \tag{2.4.3}$$

其中 $m(G_0)$ 是 G_0 的面积, $R_{G_0}(u)$ 是截断误差. 以之代到 (2.4.2), 舍去截断误

差即得点 P_0 的差分方程:

$$(2.4.4)\qquad -\sum_i \frac{\overline{q_iq_{i+1}}}{\overline{P_0P_{i+1}}}(u_{P_{i+1}}-u_{P_0})=\iint\limits_{G_0} f\mathrm{d}x\mathrm{d}y=m(G_0)\cdot\varphi_0,$$

$$\varphi_0=\frac{1}{m(G_0)}\iint\limits_{G_0} f\mathrm{d}x\mathrm{d}y.$$

其次建立界点的差分方程. 如图 4.2(b), 设 P_0 是界点, 相应的对偶单元为由 $P_0m_1q_1q_2q_3m_4P_0$ 围成的多边形 G_0. 若在 Γ 上给的是第一边值条件 $(2.3.1)_1$, 则令

$$(2.4.5)\qquad u_{P_0}=\alpha(P_0).$$

若给的是第二或第三边值条件, 例如

$$(2.4.6)\qquad \frac{\partial u}{\partial \boldsymbol{n}}+ku|_\Gamma=\gamma,$$

($k\equiv 0$ 就是第二边值条件) 则需补充一个方程. 如图 4.2(b), 此时和 (2.4.2) 类似地有

$$(2.4.7)\qquad \begin{aligned}&-\left[\int_{\overline{m_4P_0}}\frac{\partial u}{\partial \boldsymbol{n}}\mathrm{d}s+\int_{\overline{P_0m_1}}\frac{\partial u}{\partial \boldsymbol{n}}\mathrm{d}s+\int_{\overline{m_1q_1}}\frac{\partial u}{\partial \boldsymbol{n}}\mathrm{d}s+\right.\\&\left.\int_{\overline{q_1q_2}}\frac{\partial u}{\partial \boldsymbol{n}}\mathrm{d}s+\int_{\overline{q_2q_3}}\frac{\partial u}{\partial \boldsymbol{n}}\mathrm{d}s+\int_{\overline{q_3m_4}}\frac{\partial u}{\partial \boldsymbol{n}}\mathrm{d}s\right]\\&=\iint\limits_{G_0} f\mathrm{d}x\mathrm{d}y.\end{aligned}$$

上式左端方括号内后四项仿照公式 (2.4.3) 的方法离散化, 例如

$$\int_{\overline{m_1q_1}}\frac{\partial u}{\partial \boldsymbol{n}}\mathrm{d}s\approx\frac{\overline{m_1q_1}}{\overline{P_0P_1}}(u_1-u_0),\int_{\overline{q_1q_2}}\frac{\partial u}{\partial \boldsymbol{n}}\mathrm{d}s\approx\frac{\overline{q_1q_2}}{\overline{P_0P_2}}(u_2-u_0).$$

(2.4.7) 左端前两项是沿外边界的积分, 利用条件 (2.4.6) 消去法向导数, 得到仅含 u 的积分. 假定 k,γ 是常数, u 用三角单元边上的线性函数去逼近, 则可利用梯形公式计算相应的积分, 于是得

$$(2.4.8)_1\qquad \begin{aligned}\int_{\overline{m_4P_0}}\frac{\partial u}{\partial \boldsymbol{n}}\mathrm{d}s&=\int_{\overline{m_4P_0}}(\gamma-ku)\mathrm{d}s\\&=\overline{m_4P_0}(\gamma-\frac{1}{2}k(u_{P_0}+u_{m_4}))\\&=\frac{1}{2}\overline{P_0P_4}\left(\gamma-\frac{1}{4}k(3u_{P_0}+u_{P_4})\right),\end{aligned}$$

同理

$$(2.4.8)_2 \qquad \int_{\overline{P_0 m_1}} \frac{\partial u}{\partial \boldsymbol{n}} \mathrm{d}s = \frac{1}{2}\overline{P_0 P_1}\left(\gamma - \frac{1}{4}k(3u_{P_0} + u_{P_1})\right).$$

将这些公式代到 (2.4.7), 就得到界点的差分方程. 显然所有内点、界点的差分方程组成一个封闭的线性代数方程组, 其系数矩阵是对称的稀疏矩阵.

从前面的推导看出, 用有限体积法构造三角网差分格式和矩形网情形完全类似, 而边值条件的处理则更方便灵活, 特别是第二、第三边值条件, 其处理方法跟内点没有实质区别. 还应指出, 方程 (2.4.2) 是 (2.4.1) 的积分形式, 表示某一物理量在单元 G_0 守恒, 而差分方程 (2.4.4) 则是守恒律的离散形式.

例 4.1 五点差分格式

在矩形网上用同向对角线将每一矩形单元分成两个直角三角形, 得到 "直角三角剖分", 对偶单元是矩形, 由此导出的差分格式恰是五点格式.

例 4.2 正三角网上的差分格式

图 4.3 是一个正三角网, 每个三角形的边长为 h. 取一内节点 P_0, 设与之相邻的六个节点为 $P_1, P_2, P_3, P_4, P_5, P_6$. 过 $\overline{P_0P_i}$ 的中点 m_i 作 $\overline{P_0P_i}$ 的中垂线, 依次交于 $\Delta P_0P_iP_{i+1}$ 的外心 q_i (这时与重心重合). 正六边形 $q_1q_2\cdots q_6$ 围成的多边形域 G_0 是围绕 P_0 的对偶单元. 显然

$$\overline{P_0P_i} = h, \quad \overline{q_iq_{i+1}} = h/\sqrt{3}.$$

G_0 的面积 $m(G_0) = \sqrt{3}h^2/2$. 因此差分格式 (2.4.4) 为

$$(2.4.9) \qquad -\frac{2}{3h^2}\left(\sum_{i=1}^{6} u_i - 6u_0\right) = \frac{2}{\sqrt{3}h^2}\iint_{G_0} f(x,y)\mathrm{d}x\mathrm{d}y.$$

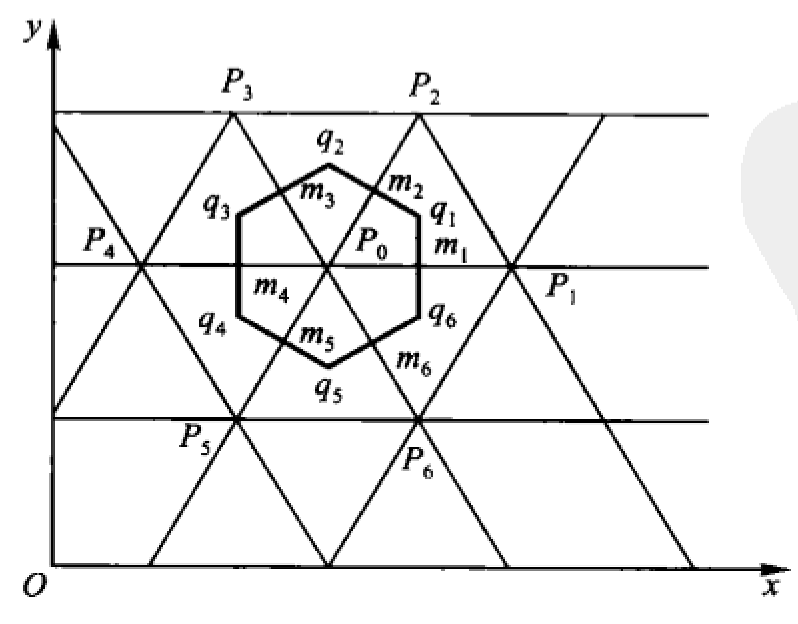


图 4.3

例 4.3 正六边形网上的差分格式

图 4.4 是正六边形网, 边长为 h. 设 P_0 是任一内节点, 与之相邻的三个节点为 P_1,P_2,P_3. 过 $\overline{P_0P_i}$ 的中点作中垂线, 彼此相交于六边形的中心 q_1,q_2,q_3, 三角形 $q_1q_2q_3$ 是围绕 P_0 的对偶单元. 显然 $\overline{P_0P_i}=h$, $\overline{q_iq_{i+1}}=\sqrt{3}h$, 对偶单元 $\Delta q_1q_2q_3$ 的面积为 $3\sqrt{3}h^2/4$. 因此差分格式为

$$-\frac{4}{3h^2}\left(\sum_{i=1}^{3}u_i-3u_0\right)=\frac{4}{3\sqrt{3}h^2}\iint\limits_{\Delta q_1q_2q_3}f(x,y)\mathrm{d}x\mathrm{d}y. \tag{2.4.10}$$

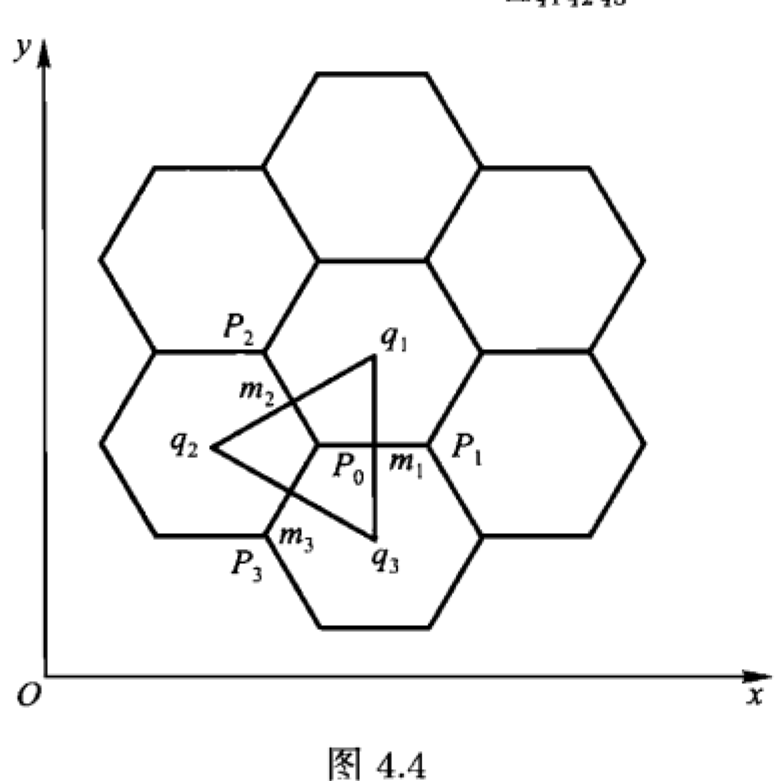


图 4.4

习 题

1. 试证三角网的差分格式 (第一或第三边值条件) 的系数矩阵对称.
2. 构造逼近 (2.3.21) 的三角网的差分格式.
3. 求出差分格式 (2.4.9) 和 (2.4.10) 的截断误差的阶 (分别为 $O(h^2)$ 和 $O(h)$).

*§5 极值定理和敛速估计

从 §1 知道, 为了得到差分解的收敛性、敛速估计及其稳定性, 需要对差分解作某种先验估计, 极值定理是作这类估计的常用方法.

5.1 差分方程

考虑二阶椭圆偏微分方程第一边值问题:

$$\begin{cases}-(Au_x)_x-(Bu_y)_y+Cu_x+Du_y+Eu=F, & (x,y)\in G.\\ u|_{\Gamma}=\alpha\end{cases} \tag{2.5.1}$$

其中 $A(x,y)$, $B(x,y)$ 属于 $C^1(\overline{G})$, $C(x,y)$, $D(x,y)$, $E(x,y)$ 和 $F(x,y)$ 属于 $C(\overline{G})$, $\alpha \in C(\Gamma)$, 且 $A(x,y) \geqslant A_{\min} > 0$, $B(x,y) \geqslant B_{\min} > 0$, $E \geqslant 0$.

如 §3 构造矩形网, h_1 和 h_2 分别为沿 x 和 y 方向的步长. 用 G_h 表示网格内点集合, Γ_h 表示网格界点集合, $\overline{G}_h = G_h \bigcup \Gamma_h$. 本节总假定 G_h 是连通的, 就是说, 对任意两节点 $\overline{P}$, $\overline{\overline{P}} \in G_h$, 必有一串节点 $P_i \in G_h$ $(i=1,2,\cdots,m-1)$, 可与 $\overline{P}, \overline{\overline{P}}$ 排成下列顺序:

$$\overline{P}, P_1, P_2, \cdots, P_{m-1}, \overline{\overline{P}},$$

使前后两点为相邻点.

对于正则内点 (x_i, y_j), 用如下的差分方程逼近 (2.5.1):

$$(2.5.2)\quad -[A_{i-\frac{1}{2},\,j}(u_{ij})_{\overline{x}}]_x - [B_{i,\,j-\frac{1}{2}}(u_{ij})_{\overline{y}}]_y + C_{ij}(u_{ij})_{\widehat{x}} + D_{ij}(u_{ij})_{\widehat{y}} + E_{ij}u_{ij} = F_{ij},$$

其中

$$A_{i-\frac{1}{2},j} = A(x_{i-\frac{1}{2}}, y_j) = A\left(\left(i-\frac{1}{2}\right)h_1, jh_2\right),$$
$$B_{i,j-\frac{1}{2}} = B(x_i, y_{j-\frac{1}{2}}) = B\left(ih_1, \left(j-\frac{1}{2}\right)h_2\right),$$

而

$$(2.5.3)\quad \begin{cases} (u_{ij})_{\overline{x}} = \dfrac{u_{ij}-u_{i-1,j}}{h_1}, (u_{ij})_{\overline{y}} = \dfrac{u_{ij}-u_{i,j-1}}{h_2}, \\ (u_{ij})_x = \dfrac{u_{i+1,j}-u_{i,j}}{h_1}, (u_{ij})_y = \dfrac{u_{i,j+1}-u_{i,j}}{h_2}, \\ (u_{ij})_{\widehat{x}} = \dfrac{u_{i+1,j}-u_{i-1,j}}{2h_1}, (u_{ij})_{\widehat{y}} = \dfrac{u_{i,j+1}-u_{i,j-1}}{2h_2}. \end{cases}$$

显然, 截断误差的阶为 $O(h_1^2+h_2^2)$. 方程 (2.5.2) 可改写成为:

$$(2.5.4)\quad a_{ij}u_{ij} - (a_{i-1,j}u_{i-1,j} + a_{i,j-1}u_{i,j-1} + a_{i+1,j}u_{i+1,j} + a_{i,j+1}u_{i,j+1}) = F_{ij},$$

其中

$$(2.5.5)\quad \begin{cases} a_{i-1,j} = h_1^{-2}\left(A_{i-\frac{1}{2},j} - \dfrac{h_1}{2}C_{ij}\right), \\ a_{i,j-1} = h_2^{-2}\left(B_{i,j-\frac{1}{2}} - \dfrac{h_2}{2}D_{ij}\right), \\ a_{i+1,j} = h_1^{-2}\left(A_{i+\frac{1}{2},j} + \dfrac{h_1}{2}C_{ij}\right), \\ a_{i,j+1} = h_2^{-2}\left(B_{i,j+\frac{1}{2}} + \dfrac{h_2}{2}D_{ij}\right), \\ a_{ij} = h_1^{-2}\left(A_{i+\frac{1}{2},j} + A_{i-\frac{1}{2},j}\right) + h_2^{-2}\left(B_{i,j+\frac{1}{2}} + B_{i,j-\frac{1}{2}}\right) + E_{ij}. \end{cases}$$

由系数 A, B 的假设条件, 只要 h_1 和 h_2 充分小, 则 $a_{i-1,j}, a_{i,j-1}, a_{i+1,j}, a_{i,j+1}$ 和 a_{ij} 均大于 0, 且

$$a_{ij} - (a_{i-1,j} + a_{i,j-1} + a_{i+1,j} + a_{i,j+1}) = E_{ij} \geqslant 0. \tag{2.5.6}$$

对于非正则内点, 则按节 3.2 的方法建立一不等距差分方程. 例如设 (x_i, y_j) 为图 3.5 中之点 "0", 则用

$$\left[A_{i-\frac{1}{2},j}(u_{ij})_{\overline{x}}\right]_x = \frac{1}{\overline{h_1}}\left[A_{i+\frac{1}{2},j}\frac{u_{i+1,j} - u_{ij}}{h_1} - A_{i-\frac{1}{2},j}\frac{u_{i,j} - u_{i-1,j}}{h_1^-}\right]$$

和

$$(u_{ij})_{\widehat{x}} = \frac{u_{i+1,j} - u_{i-1,j}}{h_1 + h_1^-}$$

依次代替 (2.5.2) 中的相应项, 其中 $h_1^-, \overline{h_1}$ 如 (2.3.12). 此时仍可将 (2.5.2) 写成形式 (2.5.4). 只要 h_1, h_1^-, h_2, h_2^- 充分小, 则 (2.5.4) 左端之系数 $a_{i-1,j}$, $a_{i,j-1}$, $a_{i+1,j}$, $a_{i,j+1}$ 和 a_{ij} 就是正的, 且 (2.5.6) 成立. 显然, 在非正则内点, 差分逼近的阶为 $O(h_1 + h_2)$.

在研究 (2.5.4) 的极值性质之前, 我们指出它的几个简单而有用的性质. 将 (2.5.4) 改写成

$$L_h u_{ij} = a_{ij}u_{ij} - a_{i-1,j}u_{i-1,j} - a_{i,j-1}u_{i,j-1} - a_{i+1,j}u_{i+1,j} - a_{i,j+1}u_{i,j+1} = F_{ij} \tag{2.5.7}$$

将网格内点按适当次序排列, 例如从左下角网点开始, 按由左向右、由下向上的顺序排列, 得到一线性代数方程组, 其系数矩阵 $\boldsymbol{A}$ 有下列性质:

(i) $\boldsymbol{A}$ 的每行最多有五个非零元素, 所以 $\boldsymbol{A}$ 为稀疏矩阵. 在排列网点顺序时, 应尽量使非零元素 "靠近" 对角线, 这对消元法特别有利.

(ii) $\boldsymbol{A}$ 的对角元素是正的, 非对角元素是非正的. 非对角元素绝对值之和不超过对角元素, 即

$$a_{i-1,j} + a_{i,j-1} + a_{i+1,j} + a_{i,j+1} \leqslant a_{ij}. \tag{2.5.8}$$

当 (i,j) 为非正则内点时, 其四个相邻点至少有一个是界点, 比如设 $(i-1,j)$ 是界点, 则可将 (2.5.7) 中之相应项 $a_{i-1,j}u_{i-1,j}$ 移到右端, 视 (2.5.7) 左端之 $a_{i-1,j} = 0$, 此时不等式 (2.5.8) 严格小于号成立. 所以矩阵 $\boldsymbol{A}$ 对角占优, 这对于保证迭代法的收敛性是重要的.

(iii) 若方程 (2.5.1) 对称, 即 $C = D = 0$, 则矩阵 $\boldsymbol{A}$ 也对称 (设非正则内点的格式为 (2.3.13)).

上述性质在构造差分方程的解法时特别有用.

5.2 极值定理

现在讨论差分方程 (2.5.7) 的极值性质.

定理 5.1 (极值定理) 设 u_{ij} 是 $\overline{G}_h$ 上的任一网格函数. 若 $L_h u_{ij} \leqslant 0$ ($L_h u_{ij} \geqslant 0$), 对任意 $(x_i, y_j) \in G_h$, 则 u_{ij} 不可能在内点取正的极大 (负的极小), 除非 $u_{ij} \equiv$ 常数.

证明 只证明定理的第一部分, 因为第二部分是类似的. 用反证法, 设 $u_{i,j} \neq$ 常数, u_{ij} 在 G_h 中某点达到正的极大值 M. 由于 G_h 连通, 必有某一内点 (x_{i_0}, y_{j_0}), 使 $u_{i_0,j_0} = M$, 且至少有一个相邻网点, 比如 (x_{i_0-1}, y_{j_0}), 使 $u_{i_0-1,j_0} < M$. 于是

$$L_h u_{i_0,j_0} > (a_{i_0,j_0} - a_{i_0-1,j_0} - a_{i_0+1,j_0-1} - a_{i_0+1,j_0} - a_{i_0,j_0+1})M \geqslant 0,$$

与假设矛盾.

推论 5.1 差分方程 (2.5.7) 有唯一解.

证明 只需证明相应的齐问题 (边值和右端都恒等于 0) 只有平凡解. 实际上, 设 u_{ij} 是齐问题的解, 则由定理 5.1, u_{ij} 既不能在 G_h 取正的极大, 也不能取负的极小, 因此 $u_{ij} \equiv 0$.

推论 5.2 若网格函数 u_{ij} 满足

$$L_h u_{ij} \geqslant 0, \text{对任意 } (x_i, y_j) \in G_h,$$
$$u_{ij} \geqslant 0, \text{对任意 } (x_i, y_j) \in \varGamma_h,$$

则 $u_{ij} \geqslant 0$, 对任意 $(x_i, y_j) \in G_h$.

证明 由定理 5.1 直接推得.

定理 5.2 (比较定理) 设 u_{ij} 和 U_{ij} 是两个网格函数, 满足

$$|L_h u_{ij}| \leqslant L_h U_{ij}, \text{ 对任意 } (x_i, y_j) \in G_h, \tag{2.5.9$_1$}$$

$$|u_{ij}| \leqslant U_{ij}, \text{ 对任意 } (x_i, y_j) \in \varGamma_h, \tag{2.5.9$_2$}$$

则

$$|u_{ij}| \leqslant U_{ij}, \quad \text{对任意 } (x_i, y_j) \in G_h. \tag{2.5.10}$$

证明 由 $(2.5.9)_1$ 和 $(2.5.9)_2$ 可知

$$\begin{cases} L_h(U_{ij} - u_{ij}) \geqslant 0, & \text{于 } G_h, \\ U_{ij} - u_{ij} \geqslant 0, & \text{于 } \varGamma_h \end{cases}$$

和

$$\begin{cases} L_h(U_{ij}+u_{ij}) \geqslant 0, & 于\ G_h, \\ U_{ij}+u_{ij} \geqslant 0, & 于\ \Gamma_h. \end{cases}$$

由推论 5.2 便知 (2.5.10) 成立.

推论 5.3 *差分方程*

$$\begin{cases} L_h u_{ij} = 0, & 于\ G_h, \\ u_{ij} = \alpha_{ij}, & 于\ \Gamma_h \end{cases}$$

的解 u_{ij} 满足不等式

$$\max_{G_h} |u_{ij}| \leqslant \max_{\Gamma_h} |\alpha_{ij}|. \tag{2.5.11}$$

证明 设 U_{ij} 是下列问题的解:

$$L_h U_{ij} = 0, \quad 于\ G_h,$$
$$U_{ij} = |\alpha_{ij}|, \quad 于\ \Gamma_h,$$

则由定理 5.2

$$|u_{ij}| \leqslant U_{ij}, 于\ \overline{G}_h.$$

若 $U_{ij} \equiv$ 常数, 则 $U_{ij} \equiv \max\limits_{\Gamma_h} |\alpha_{ij}|$. 若 $U_{ij} \neq$ 常数, 则由定理 5.1, 函数 $U_{ij}(\geqslant 0)$ 的最大值只能在 Γ_h 达到, 因此 $U_{ij} \leqslant \max\limits_{\Gamma_h} |\alpha_{ij}|$, 从而 (2.5.11) 成立.

5.3 五点格式的敛速估计

设 $u = u(x,y)$ 是 Poisson 方程第一边值问题的解 (参看 (2.3.1), $(2.3.1)_1$), u_{ij} 是五点差分格式

$$\begin{cases} L_h u_{ij} = \varphi_{ij}, & 于\ G_h \\ u_{ij} = \alpha_{ij}, & 于\ \Gamma_h \end{cases}$$

的解, 此时 $A = B = 1$, $C = D = E = 0$. 当 (x_i, y_j) 是正则内点, 且 $h_1 = h_2$ 时, (2.5.7) 中的 $a_{ij} = 4$, $a_{i-1,j} = a_{i,j-1} = a_{i+1,j} = a_{i,j+1} = 1$. 当 (x_i, y_j) 为非正则内点时, 需按节 3.2 的方法对系数作适当修正. 令 $e_{ij} = u(x_i, y_j) - u_{ij}$, 若 $u \in C^4(\overline{G})$, 则 e_{ij} 满足

$$\begin{cases} L_h e_{ij} = R_{ij}, & 于\ G_h \\ e_{ij} = 0, & 于\ \Gamma_h. \end{cases} \tag{2.5.12}$$

截断误差

$$R_{ij} = \begin{cases} O(h_1^2 + h_2^2), & (x_i, y_j)\ 是正则内点, \\ O(h_1 + h_2), & (x_i, y_j)\ 是非正则内点. \end{cases}$$

令 $h=\sqrt{h_1^2+h_2^2}$, 可设

$$|R_{ij}|\leqslant Kh,$$

其中 K 是常数.

现在估计误差 e_{ij}. 不妨设 $(0,0)\in G, R$ 是以 $(0,0)$ 为心且包含 G 在内部的最小圆域的半径. 令

$$E_{ij}=\frac{Kh}{4}(R^2-x_i^2-y_j^2).$$

显然, E_{ij} 于 $\overline{G_h}$ 非负. 又 x_i^2 关于 x 方向的二阶中心差商等于 2, y_j^2 关于 y 方向的二阶中心差商也等于 2, 故当 (x_i,y_j) 是正则内点时, $L_hE_{ij}=Kh$. 而当 (x_i,y_j) 为非正则内点时, L_h 中仍出现二阶差商 (但不一定是中心差商), 此时仍有 $L_hE_{ij}=Kh$. 这样, E_{ij} 满足

$$\begin{cases} L_hE_{ij}=Kh, & (x_i,y_j)\in G_h, \\ E_{ij}\geqslant 0, & (x_i,y_j)\in \Gamma_h. \end{cases} \tag{2.5.13}$$

将定理 5.2 用于 (2.5.12) 和 (2.5.13), 便得

$$|e_{ij}|\leqslant E_{ij}=\frac{Kh}{4}(R^2-x_i^2-y_j^2),$$

从而

$$\max_{G_h}|e_{ij}|\leqslant\frac{KR^2h}{4}. \tag{2.5.14}$$

足见若 $u(x,y)\in C^4(\overline{G})$, 则差分解 u_{ij} 一致收敛到 u, 且有敛速估计 (2.5.14)

注 极值定理在证明椭圆差分方程的稳定性及差分解的误差估计中起重要作用. 在下列习题 1,2 中给出极值定理的其他形式, 它们将在第三章研究隐式差分格式理论中得到应用.

习　　题

1. 设 $\overline{I}_h=\{x_i:i=0,1,\cdots,N,x_0<x_1<\cdots<x_N\}$, y_i 是 $\overline{I}_h$ 上的网格函数. 又

$$ly_i=-(a_iy_{i-1}-b_iy_i+c_iy_{i+1})+q_iy_i, i=1,2,\cdots,N-1,$$

其中 a_i, b_i, c_i 恒正, q_i 非负, 且 $a_i+c_i\leqslant b_i$. 证明当 $ly_i\leqslant 0$ ($ly_i\geqslant 0$) 时, y_i 不能在内点取正的极大 (负的极小), 除非 y_i 等于常数.

2. 在题 1 中, 若设 $d_i=b_i-a_i-c_i+q_i>0$ $(i=1,2,\cdots,N-1)$, 则差分方程

$$\begin{cases} ly_i=\varphi_i, i=1,2,\cdots,N-1, \\ y_0=y_N=0 \end{cases}$$

的解满足

$$\max_i|y_i|\leqslant\max_i\frac{|\varphi_i|}{d_i}.$$

3. 利用题 2 估计差分方程 (2.2.9), (2.2.10) 解的收敛阶 (假定 $r=0, q\geqslant q_0>0, h_i\equiv h$).

§6 迭 代 法

在本章最后三节, 我们介绍差分方程的解法, 包括逐次超松弛法, 共轭梯度法和交替方向法. 这些方法也可用于第六章要介绍的 Galerkin 有限元法. 尽管读者可能在一般线性代数计算教程学过这些方法, 但由偏微分方程数值解法导出的离散化方程有其特点, 所以仍有必要知道如何利用方程的特性选择适当解法. 其实, 现行线性代数计算的许多方法也是从偏微分方程数值解中提出的. 我们在前节已经指出, 用差分法 (或有限元法) 导出的线性代数方程的系数矩阵一般具有性质: 大型稀疏, 对角占优, 若原方程对称正定, 则相应差分方程的系数矩阵也对称正定. 现在还要指出, 如此的矩阵一般还是病态的, 这正是求解这类方程的困难所在.

设 $\boldsymbol{A}$ 是一非奇异 N 阶矩阵, $\lambda_1, \lambda_2, \cdots, \lambda_N$ 是 $\boldsymbol{A}$ 的特征值, 其中 $|\lambda_1| \leqslant |\lambda_2| \leqslant \cdots \leqslant |\lambda_N|$. 通常称 $\|\boldsymbol{A}\| \cdot \|\boldsymbol{A}^{-1}\|$ 为矩阵 $\boldsymbol{A}$ 的条件数, 而称 $|\lambda_N|/|\lambda_1|$ 为 $\boldsymbol{A}$ 的谱条件数. 当 $\boldsymbol{A}$ 是对称矩阵时谱条件数和条件数相同. 在一般情形, $|\lambda_N|/|\lambda_1| \leqslant \|A\| \cdot \|A^{-1}\|$. 往后记 $\mathrm{Cond}(\boldsymbol{A}) = |\lambda_N|/|\lambda_1|$ 并称之为**条件数**. 显然恒有 $\mathrm{Cond}(\boldsymbol{A}) \geqslant 1$. 若 $\mathrm{Cond}(\boldsymbol{A}) >> 1$, 则说 A 是**病态矩阵**. 我们知道, 当 $\boldsymbol{A}$ 是病态时, 求解方程组 $\boldsymbol{A}u = f$ 会遇到很多困难. 下面我们举例说明, 由椭圆型方程导出的差分方程一般都是病态的.

例 6.1 考虑两点边值问题:

$$
\begin{aligned}
&Lu = -\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} + qu = f, \quad 0 < x < 1, \\
&u(0) = u(1) = 0,
\end{aligned} \tag{2.6.1}
$$

$q \geqslant 0$ 是常数. 取步长 $h = 1/N$, 用二阶中心差商替代二阶微商, 得差分方程:

$$
\begin{aligned}
&-u_{j-1} + (2 + h^2 q)u_j - u_{j+1} = h^2 f_j, \quad j = 1, 2, \cdots, N-1, \\
&u_0 = u_N = 0,
\end{aligned} \tag{2.6.2}
$$

相应的特征问题为

$$
\begin{aligned}
&L_h u_j = (l_h + (2 + h^2 q)I)u_j = \lambda u_j, \quad j = 1, 2, \cdots, N-1, \\
&u_0 = u_N = 0,
\end{aligned} \tag{2.6.3}
$$

其中 l_h 是由下式确定的差分算子:

$$
\begin{aligned}
&l_h u_j = -(u_{j-1} + u_{j+1}), \quad j = 1, 2, \cdots, N-1, \\
&u_0 = u_N = 0.
\end{aligned} \tag{2.6.4}
$$

先求 $l_h u_j$ 的特征值和特征向量. 考虑

(2.6.5)
$$l_h u_j = -(u_{j-1} + u_{j+1}) = \mu u_j,$$
$$u_0 = u_N = 0$$

以 $u_j = z^j$ 代入, 为使 (2.6.5) 有非零解, 必须

(2.6.6)
$$\mu = \mu_i = -2\cos i\pi h, \quad i = 1, 2, \cdots, N-1.$$

相应的解向量

(2.6.7)
$$\boldsymbol{u}^i = \{\sin ij\pi h\}_{j=1}^{N-1}, \quad i = 1, 2, \cdots, N-1.$$

μ_i 是 (2.6.5) 的特征值, $\boldsymbol{u}^i$ 是相应的特征向量. 由此可知 L_h ((2.6.3)) 的特征值为

$$\lambda_i = -2\cos i\pi h + 2 + h^2 q = 4\sin^2\frac{i\pi h}{2} + qh^2,$$

特征向量仍为 $\boldsymbol{u}^i$ ((2.6.7)). 最小和最大特征值分别为

$$\lambda_1 = 4\sin^2\frac{\pi h}{2} + qh^2 \sim (\pi^2 + q)h^2, \quad h \to 0$$

和

$$\lambda_{N-1} = 4\sin^2\frac{(N-1)\pi h}{2} + qh^2 \sim 4, \quad h \to 0.$$

条件数

$$\frac{\lambda_{N-1}}{\lambda_1} \sim \frac{4}{\pi^2 + q}h^{-2}, \quad h \to 0.$$

所以当 h 充分小时, 系数矩阵是病态的.

例 6.2 考虑二维模型问题:

(2.6.8)
$$-\Delta u = -(u_{xx} + u_{yy}) = f,$$
$$(x, y) \in G = (0,1) \times (0,1) = \{0 < x < 1; 0 < y < 1\}, u|_\Gamma = 0,$$

其中 $\Gamma = \partial G$ 是 G 得边界. 取步长 $h = 1/N$, 作正方形网格, 则逼近 (2.6.8) 的五点差分格式为

(2.6.9)
$$-u_{i-1,j} - u_{i,j-1} + 4u_{ij} - u_{i+1,j} - u_{i,j+1} = h^2 f_{ij},$$
$$u_{i0} = u_{0j} = u_{iN} = u_{Nj} = 0, \quad i, j = 1, 2, \cdots, N-1.$$

用分离变量法, 取

(2.6.10)
$$u_{ij}^{(r,s)} = \sin\frac{ri\pi}{N}\sin\frac{sj\pi}{N}, \quad r, s = 1, 2, \cdots, N-1,$$

则

$$-u_{i-1,j}^{(r,s)} - u_{i,j-1}^{(r,s)} + 4u_{ij}^{(r,s)} - u_{i+1,j}^{(r,s)} - u_{i,j+1}^{(r,s)} = 4\left(\sin^2\frac{r\pi}{2N} + \sin^2\frac{s\pi}{2N}\right)u_{ij}^{(r,s)},$$

足见由 (2.6.9) 左端定义的差分算子 (或系数矩阵) 的特征值是

$$\lambda^{(r,s)} = 4\left(\sin^2\frac{r\pi}{2N} + \sin^2\frac{s\pi}{2N}\right), \quad r,s = 1,2,\cdots,N-1, \tag{2.6.11}$$

相应的特征向量 (由 (2.6.10) 定义) 是

$$\boldsymbol{u}^{(r,s)} = \left\{\sin\frac{ri\pi}{N}\sin\frac{sj\pi}{N}\right\}_{i,j=1}^{N-1}. \tag{2.6.12}$$

最小特征值

$$\lambda^{(1,1)} = 8\sin^2\frac{\pi h}{2} \sim 2\pi^2 h^2, \quad h \to 0;$$

最大特征值

$$\lambda^{(N-1,N-1)} = 8\sin^2\left(\frac{1}{2} - \frac{h}{2}\right)\pi = 8\cos^2\frac{\pi h}{2} \sim 8, \quad h \to 0;$$

条件数

$$\lambda^{(N-1,N-1)}/\lambda^{(1,1)} \sim \frac{4}{\pi^2}h^{-2}, \quad h \to 0.$$

故对充分小的 h, (2.1.9) 的系数矩阵是病态的.

五点差分格式的系数矩阵还有一个重要性质: 可排成分块三对角阵. 实际上, 将网格点 (x_i, y_j) 按 $i+j$ 的次序由小到大由下到上排列, 如图 6.1,

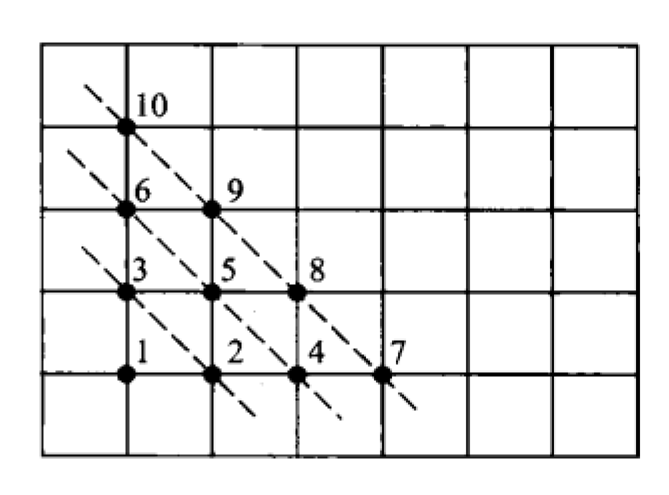


图 6.1

则五点差分格式的系数矩阵形如

$$A = \begin{bmatrix} \boldsymbol{D}_1 & \boldsymbol{C}_1 & & & \boldsymbol{0} \\ \boldsymbol{B}_1 & \boldsymbol{D}_2 & \boldsymbol{C}_2 & & \\ & \boldsymbol{B}_2 & \ddots & \ddots & \\ \boldsymbol{0} & & \ddots & \ddots & \boldsymbol{C}_{m-1} \\ & & & \boldsymbol{B}_{m-1} & \boldsymbol{D}_m \end{bmatrix}$$

其中 $\boldsymbol{D}_1, \cdots, \boldsymbol{D}_m$ 是不同阶的对角方阵. 文献上把这类矩阵归结为**2-循环矩阵**或具有性质 (A) 的矩阵 (参看 [4],[30],[32]).

6.1 一般迭代法

将线代数方程组

$$\boldsymbol{Au} = \boldsymbol{b} \tag{2.6.13}$$

改写成迭代形式:

$$\boldsymbol{u} = \boldsymbol{Tu} + \boldsymbol{d}. \tag{2.6.14}$$

解 (2.6.13) 的一般迭代程序为

$$\boldsymbol{u}^{(k+1)} = \boldsymbol{Tu}^{(k)} + \boldsymbol{d}, k = 0, 1, \cdots, \tag{2.6.15}$$

由第一章引理 1.1 推出, 它收敛的充要条件是矩阵 $\boldsymbol{T}$ 的谱半径 $\rho(\boldsymbol{T}) < 1$. 若 $\boldsymbol{T}$ 为对称矩阵,则还有敛速估计:

$$\|\boldsymbol{u}^{(k)} - \boldsymbol{u}\| \leqslant \rho^k(\boldsymbol{T})\|\boldsymbol{u}^{(0)} - \boldsymbol{u}\|,$$

其中 $\|\cdot\|$ 是欧氏模. 为使 k 次迭代误差不超过初始误差的 ε 倍, 应取 k 满足 $\rho^k(\boldsymbol{T}) \leqslant \varepsilon$ 或

$$k \geqslant |\ln \varepsilon|/(-\ln \rho(\boldsymbol{T})). \tag{2.6.16}$$

定义敛速为

$$R(\boldsymbol{T}) = -\ln \rho(\boldsymbol{T}). \tag{2.6.17}$$

设 N_0 是每步迭代所需的工作量, 则 k 次迭代的总工作量

$$M \approx N_0 |\ln \varepsilon| \big/ R(\boldsymbol{T}) \tag{2.6.18}$$

现在介绍一般的同步迭代. 假定 $\boldsymbol{A}$ 的对角元素 $a_{ii} \neq 0 (i = 1, 2, \cdots, N)$, $\boldsymbol{D} = \operatorname{diag}(a_{11}, \cdots, a_{NN})$, $\boldsymbol{B} = \boldsymbol{D} - \boldsymbol{A}$, 则方程 (2.6.13) 即 $\boldsymbol{Du} - \boldsymbol{Bu} = \boldsymbol{b}$, 两端乘以 $\boldsymbol{D}^{-1} = \operatorname{diag}(a_{11}^{-1}, \cdots, a_{NN}^{-1})$, 得

$$\boldsymbol{u} = \boldsymbol{D}^{-1}\boldsymbol{Bu} + \boldsymbol{D}^{-1}\boldsymbol{b} = (\boldsymbol{I} - \boldsymbol{D}^{-1}\boldsymbol{A})\boldsymbol{u} + \boldsymbol{D}^{-1}\boldsymbol{b}.$$

熟知的 Jacobi 迭代为

$$\boldsymbol{u}^{(k+1)} = (\boldsymbol{I} - \boldsymbol{D}^{-1}\boldsymbol{A})\boldsymbol{u}^{(k)} + \boldsymbol{D}^{-1}\boldsymbol{b}. \tag{2.6.19}$$

按分量展开,

$$u_i^{(k+1)} = -\frac{1}{a_{ii}}\sum_{j\neq i} a_{ij}u_j^{(k)} + \frac{1}{a_{ii}}b_i, \quad i = 1, 2, \cdots, N.$$

将它用于 (2.6.2) ($q = 0$), 则

$$\begin{aligned} u_i^{(k+1)} &= \frac{1}{2}(u_{i-1}^{(k)} + u_{i+1}^{(k)}) + \frac{h^2}{2}f_i, \quad i = 1, 2, \cdots, N-1, \\ u_0^{(k+1)} &= u_N^{(k+1)} = 0. \end{aligned}$$

迭代矩阵 $\boldsymbol{T} = -\frac{1}{2}\boldsymbol{l}_h$ (矩阵 $\boldsymbol{l}_h$ 由 (2.6.4) 定义) 的特征值为

$$\lambda_i = \cos i\pi h, \quad i = 1, 2, \cdots, N-1, h = N^{-1},$$

可见

$$\begin{aligned} \rho(\boldsymbol{T}) &= \cos\pi h \sim 1 - \frac{1}{2}\pi^2h^2, \quad h\to 0, \\ \boldsymbol{R}(T) &= -\ln\rho(\boldsymbol{T}) \sim \frac{1}{2}\pi^2h^2, \quad h \to 0. \end{aligned} \tag{2.6.20}$$

引进松弛因子 $\omega \neq 0$ (实参数), 将方程 (2.6.13) 写成等价形式: $\boldsymbol{u} = \boldsymbol{u} - \omega(\boldsymbol{A}\boldsymbol{u} - \boldsymbol{b})$ 就得到 **Richardson** 同步迭代:

$$\boldsymbol{u}^{(k+1)} = \boldsymbol{u}^{(k)} - \omega(\boldsymbol{A}\boldsymbol{u}^{(k)} - \boldsymbol{b}), \quad k = 0, 1, \cdots, \tag{2.6.21}$$

迭代矩阵 $\boldsymbol{T} = \boldsymbol{I} - \omega\boldsymbol{A}$, 其特征值

$$\lambda_i^{\boldsymbol{T}} = 1 - \omega\lambda_i, \quad i = 1, 2, \cdots, N$$

λ_i 是 $\boldsymbol{A}$ 的特征值. 假定 $\boldsymbol{A}$ 是对称正定矩阵,α 和 β 是其特征值的下界和上界:

$$0 < \alpha \leqslant \lambda_i \leqslant \beta, \quad i = 1, 2, \cdots, N.$$

为使 $\rho(\boldsymbol{T}) = \max_i |1 - \omega\lambda_i| < 1$ 必须且只须

$$0 < \omega < \frac{2}{\beta}.$$

选最佳松弛因子 ω_{opt} 使

$$\rho_{\omega_{opt}} = \min_\omega \max_i |1 - \omega\lambda_i|.$$

显然 $1-\omega x$ 的最大和最小值在端点达到, 使 $\max\limits_i|1-\omega\lambda_i|$ 最小的 ω_{opt} 满足

$$1-\omega_{opt}\alpha=-(1-\omega_{opt}\beta),$$

(2.6.22)
$$\omega_{opt}=\frac{2}{\alpha+\beta}.$$

相应的

(2.6.23)
$$\rho(\boldsymbol{T})=\rho(\boldsymbol{I}-\omega_{opt}\boldsymbol{A})=\frac{\beta-\alpha}{\beta+\alpha}.$$

将 Richardson 迭代用于五点格式 (2.6.9), 因 $\boldsymbol{A}$ 的特征值

$$\lambda^{(i,j)}=4\sin^2\frac{i\pi h}{2}+4\sin^2\frac{j\pi h}{2},$$
$$\alpha=\lambda^{(1,1)}=8\sin^2\frac{\pi h}{2},\ \beta=\lambda^{(N-1,N-1)}=8\cos^2\frac{\pi h}{2},$$

故 $\omega_{opt}=\dfrac{1}{4}$, 相应的迭代为

(2.6.24)
$$\begin{aligned}u_{ij}^{(k+1)}&=u_{ij}^{(k)}-\frac{1}{4}(-u_{i-1,j}^{(k)}-u_{i,j-1}^{(k)}+4u_{ij}^{(k)}-u_{i+1,j}^{(k)}-u_{i,j+1}^{(k)}-h^2f_{ij})\\&=\frac{1}{4}(u_{i-1,j}^{(k)}+u_{i,j-1}^{(k)}+u_{i+1,j}^{(k)}+u_{i,j+1}^{(k)}+h^2f_{ij}),\quad k=0,1,\cdots,\end{aligned}$$

这是解 (2.6.9) 的 **Jacobi** 迭代, 迭代矩阵 $\boldsymbol{T}$ 的特征值 $\lambda_{ij}^{\boldsymbol{T}}=1-\dfrac{1}{4}\lambda_{ij}$, 谱半径

(2.6.25)
$$\rho(\boldsymbol{T})=\cos\pi h\sim1-\frac{\pi^2h^2}{2},\quad h\to0.$$

收敛速度

(2.6.26)
$$R(\boldsymbol{T})=-\ln\rho(\boldsymbol{T})\sim\frac{\pi^2h^2}{2},\quad h\to0.$$

6.2 SOR 法 (逐次超松弛法)

将二阶椭圆方程的五点差分格式 (2.5.4) 改写成:

(2.6.27)
$$L_hu_{i,j}=a_{ij}u_{ij}-a_{i-1,j}u_{i-1,j}-a_{i,j-1}u_{i,j-1}-a_{i+1,j}u_{i+1,j}-a_{i,j+1}u_{i,j+1}=f_{ij}.$$

如例 6.2, 将网点按下标之和 $i+j=l$ 递增的次序排列, 则知其系数矩阵 $\boldsymbol{A}$ 具有性质 (A). 格式 (2.6.27) 的系数都是非负的: $a_{ij}\geqslant0$, $a_{i-1,j}\geqslant0$, $a_{i,j-1}\geqslant0$, $a_{i+1,j}\geqslant0$, $a_{i,j+1}\geqslant0$, 且对角占优:

$$0<a_{i-1,j}+a_{i,j-1}+a_{i+1,j}+a_{i,j+1}\leqslant a_{ij}.$$

(当节点 (x_i, y_j) 有一邻点是界点时第二个不等号严格成立). 解 (2.6.27) 的 **SOR 法**为

$$
(2.6.28)\qquad \begin{aligned} u_{ij}^{(k+1)} = u_{ij}^{(k)} - \frac{\omega}{a_{ij}}(a_{ij}u_{ij}^{(k)} - a_{i-1,j}u_{i-1,j}^{(k+1)} - a_{i,j-1}u_{i,j-1}^{(k+1)} \\ - a_{i+1,j}u_{i+1,j}^{(k)} - a_{i,j+1}u_{i,j+1}^{(k)} - f_{ij}), \quad i,j,k = 0,1,\cdots, \end{aligned}
$$

其中 $\omega \neq 0$ 是松弛因子. 设 $\boldsymbol{D}$ 是 $\boldsymbol{A}$ 的对角元矩阵, $-\boldsymbol{L}$ 和 $-\boldsymbol{R}$ 分别为 $\boldsymbol{A}$ 的下三角形和上三角形矩阵, $\boldsymbol{u}$ 是以 $\{u_{ij}\}$ 为分量的向量 (按前述次序排列), 则可将 (2.6.28) 写成向量形式:

$$
(2.6.29)\qquad \begin{aligned} \boldsymbol{u}^{(k+1)} &= \boldsymbol{A}_\omega \boldsymbol{u}^{(k)} + \omega(\boldsymbol{D} - \omega\boldsymbol{L})^{-1}\boldsymbol{f}, \\ \boldsymbol{A}_\omega &= (\boldsymbol{D} - \omega\boldsymbol{L})^{-1}(\omega\boldsymbol{R} + (1-\omega)\boldsymbol{D}). \end{aligned}
$$

为便于比较, 我们写下 Jacobi 迭代:

$$
(2.6.30)\qquad \begin{aligned} u_{ij}^{(k+1)} = u_{ij}^{(k)} - \frac{1}{a_{ij}}(a_{ij}u_{ij}^{(k)} - a_{i-1,j}u_{i-1,j}^{(k)} - a_{i,j-1}u_{i,j-1}^{(k)} \\ - a_{i+1,j}u_{i+1,j}^{(k)} - a_{i,j+1}u_{i,j+1}^{(k)} - f_{ij}), \end{aligned}
$$

其迭代矩阵 $\boldsymbol{B} = \boldsymbol{I} - \boldsymbol{D}^{-1}\boldsymbol{A}$.

由于矩阵 $\boldsymbol{A}$ 具性质 (A), 故 $\boldsymbol{A}_\omega$ 的特征值 λ_ω 和 Jacobi 迭代矩阵 $\boldsymbol{B}$ 的特征值 μ 之间有关系 (参看 [4], [30], [32]):

$$
\lambda_\omega - \omega\mu\lambda_\omega^{\frac{1}{2}} + \omega - 1 = 0.
$$

假定 $\boldsymbol{B}$ 的特征值都是实的, 则可推出 $|\lambda_\omega| < 1$ 的充要条件是:

$$
(2.6.31)\qquad 0 < \omega < 2, \quad |\mu| < 1.
$$

也就是说 SOR 法收敛当且仅当 $\omega \in (0,2)$ 且 Jacobi 迭代 (2.6.30) 收敛. 此外, 设 $\overline{\mu}$ 是 $\boldsymbol{B}$ 的按模最大特征值, 则**最佳松弛因子**

$$
(2.6.32)\qquad \omega_{opt} = 2/(1+\sqrt{1-\overline{\mu}^2})\ (>1).
$$

$\boldsymbol{A}_{\omega_{opt}}$ 的谱半径

$$
(2.6.33)\qquad \rho(\boldsymbol{A}_{\omega_{opt}}) = \omega_{opt} - 1 = \frac{1-\sqrt{1-\overline{\mu}^2}}{1+\sqrt{1-\overline{\mu}^2}}
$$

(参看 [4], [30], [32]).

将 SOR 法用于五点格式 (2.6.2), 因 $\boldsymbol{B}=\boldsymbol{I}-\dfrac{\boldsymbol{A}}{4}$ 的最大特征值 $\overline{\mu}=\cos\pi h$, 故

$$\rho(\boldsymbol{A}_{\omega_{opt}})=\frac{1-\sin\pi h}{1+\sin\pi h}\sim 1-2\pi h,\quad h\to 0.$$

敛速

$$R(\boldsymbol{A}_{\omega_{opt}})=-\ln\rho(\boldsymbol{A}_{\omega_{opt}})\sim 2\pi h. \tag{2.6.34}$$

由 (2.6.26), 知 Jacobi 迭代的敛速为 $O(h^2)$, 可见 SOR 法较 Jacobi 迭代有明显改进.

Gauss-Seidel 法对应 $\omega=1$, 此时 $\lambda_m^{\frac{1}{2}}=\overline{\mu}=\cos\pi h$, ($\lambda_m$ 是 $\boldsymbol{A}_1$ 的最大特征值),

$$R(\boldsymbol{A}_1)=-\ln|\lambda_m|=-2\ln\overline{\mu}\sim\pi^2h^2,$$

与 (6.6.26) 比较, 可见 Gauss-Seidel 法的敛速是 Jacobi 迭代的二倍.

注 SOR 法也可按与前述网点相反的次序进行. 还可按这两种次序交替使用, 称为对称逐次超松弛或 SSOR, 敛速比 SOR 法略有提高 (参看 [30], [32]).

习 题

1. 求 N 阶三对角阵

$$\boldsymbol{C}=\begin{bmatrix}0 & 1 & & 0\\ 1 & 0 & \ddots & \\ & \ddots & \ddots & 1\\ 0 & & 1 & 0\end{bmatrix}$$

的特征值和特征向量 (参看例 6.1).

2. 设方程组 $\boldsymbol{u}=\boldsymbol{H}\boldsymbol{u}+\boldsymbol{b}$ 满足 $\rho(\boldsymbol{H})<1$, 且 $\boldsymbol{H}$ 的特征值为实数. 证明由迭代法

$$\begin{aligned}\boldsymbol{u}^{(k+1)}&=\boldsymbol{u}^{(k)}+\omega(\boldsymbol{H}\boldsymbol{v}^{(k)}+\boldsymbol{b}-\boldsymbol{u}^{(k)}),\\ \boldsymbol{v}^{(k+1)}&=\boldsymbol{v}^{(k)}+\omega(\boldsymbol{H}\boldsymbol{u}^{(k+1)}+\boldsymbol{b}-\boldsymbol{v}^{(k)}),\end{aligned}\quad k=0,1,\cdots,$$

(对 $\omega\in(0,2)$) 确定的 $\boldsymbol{u}^{(k)}$, $\boldsymbol{v}^{(k)}$ 均收敛到方程组的解 $\boldsymbol{u}$. (提示: 原方程和联立方程 $\boldsymbol{u}=\boldsymbol{H}\boldsymbol{v}+\boldsymbol{b}$, $\boldsymbol{v}=\boldsymbol{H}\boldsymbol{u}+\boldsymbol{b}$ 等价.)

3. (实习题) 求解边值问题:

$$\begin{cases}\Delta u=2\pi^2e^{\pi(x+y)}\left(\sin\pi x\cos\pi y+\cos\pi x\sin\pi y\right),\\ (x,y)\in G=(0,1)\times(0,1),\\ u=0,(x,y)\in\partial G.\end{cases} \tag{2.6.35}$$

取步长 $h=k=1/64, 1/128$, 作五点差分格式. 用 Jacobi 迭代, Gauss-Seidel 迭代和 SOR 迭代 (取 $\varpi=\varpi_{opt}$) 求解差分方程, 以前后两次重合到小数点后四位的迭代值作为解的近似.

比较三种解法的迭代次数以及差分解 $u_h(x,y)$ $(h=1/64,1/128)$ 与精确解

$$u(x,y)=e^{\pi(x+y)}\sin\pi x\sin\pi y$$

的精度.

§7 交替方向迭代法

SOR 法虽然比 Jacobi 迭代有了高一阶的敛速, 但对应用来说 (比如在油田开发计算中) 仍然是不够的. 本节介绍一类有更高敛速的隐式迭代法 —— 交替方向迭代法.

考虑边值问题

$$\begin{aligned} &-\Delta u=f(x,y), \quad 0<x,y<1,\\ &u=0,\ \text{当}\ x=0,1\ \text{或}\ y=0,1 \end{aligned} \tag{2.7.1}$$

的五点差分格式

$$-\Delta_h u_{ij}=-\left(\frac{u_{i+1,j}-2u_{ij}+u_{i-1,j}}{h^2}+\frac{u_{i,j+1}-2u_{i,j}+u_{i,j-1}}{h^2}\right)=f_{ij}. \tag{2.7.2}$$

历史上第一个交替方向迭代法是由 Peaceman 和 Rachford 于 1955 年提出的. 先定义矩阵 $\boldsymbol{L}_1$ 和 $\boldsymbol{L}_2$. 对向量 $\boldsymbol{u}=\{u_{ij}\}$, 令

$$\begin{aligned} (\boldsymbol{L}_1\boldsymbol{u})_{ij}&=(-u_{i-1,j}+2u_{ij}-u_{i+1,j})/h^2,\\ (\boldsymbol{L}_2\boldsymbol{u})_{ij}&=(-u_{i,j-1}+2u_{ij}-u_{i,j+1})/h^2,\\ (\boldsymbol{L}_h\boldsymbol{u})_{ij}&=(\boldsymbol{L}_1\boldsymbol{u})_{ij}+(\boldsymbol{L}_2\boldsymbol{u})_{ij}, \end{aligned}$$

则可将 (2.7.2) 写成

$$\boldsymbol{L}_h\boldsymbol{u}=\boldsymbol{L}_1\boldsymbol{u}+\boldsymbol{L}_2\boldsymbol{u}=\boldsymbol{f},$$

其中 $\boldsymbol{f}=\{f_{ij}\}$. 于是 **Peaceman-Rachford 迭代** (简称 **PR 迭代**) 为:

$$\boldsymbol{u}^{(k+\frac12)}=\boldsymbol{u}^{(k)}-\tau_k(\boldsymbol{L}_1\boldsymbol{u}^{(k+\frac12)}+\boldsymbol{L}_2\boldsymbol{u}^{(k)}-\boldsymbol{f}), \tag{2.7.3$_1$}$$

$$\boldsymbol{u}^{(k+1)}=\boldsymbol{u}^{(k+\frac12)}-\tau_k(\boldsymbol{L}_1\boldsymbol{u}^{(k+\frac12)}+\boldsymbol{L}_2\boldsymbol{u}^{(k+1)}-\boldsymbol{f}), \quad k=0,1,\cdots. \tag{2.7.3$_2$}$$

其中 τ_k 是迭代参数. 按层合并, 得

$$(\boldsymbol{I}+\tau_k\boldsymbol{L}_1)\boldsymbol{u}^{(k+\frac12)}=(\boldsymbol{I}-\tau_k\boldsymbol{L}_2)\boldsymbol{u}^{(k)}+\tau_k\boldsymbol{f}, \tag{2.7.3$'_1$}$$

$$(\boldsymbol{I}+\tau_k\boldsymbol{L}_2)\boldsymbol{u}^{(k+1)}=(\boldsymbol{I}-\tau_k\boldsymbol{L}_1)\boldsymbol{u}^{(k+\frac12)}+\tau_k\boldsymbol{f}. \tag{2.7.3$'_2$}$$

由 $\boldsymbol{u}^{(k)}$ 到 $\boldsymbol{u}^{(k+1)}$ 只需交替解两个具三对角系数阵的方程 $(2.7.3)_1'$ 和 $(2.7.3)_2'$, 这不难用消元法实现 (见习题 2). 由于 $(2.7.3)_{1-2}'$ 交替沿 x 和 y 方向求解, 所以称为交替方向隐式迭代法.

现在研究方法的收敛性和迭代参数的选择. 显然差分解 $\boldsymbol{u}=\{u_{ij}\}$ 满足 $(2.7.3)_1'$, $(2.7.3)_2'$. 因此迭代误差 $\boldsymbol{e}^{(k)}=\boldsymbol{u}-\boldsymbol{u}^{(k)}$ 满足齐方程:

$$(\boldsymbol{I}+\tau_k\boldsymbol{L}_1)\boldsymbol{e}^{(k+\frac{1}{2})}=(\boldsymbol{I}-\tau_k\boldsymbol{L}_2)\boldsymbol{e}^{(k)}, \tag{2.7.4$_1$}$$

$$(\boldsymbol{I}+\tau_k\boldsymbol{L}_2)\boldsymbol{e}^{(k+1)}=(\boldsymbol{I}-\tau_k\boldsymbol{L}_1)\boldsymbol{e}^{(k+\frac{1}{2})}. \tag{2.7.4$_2$}$$

取 $\tau_k>0$, 则左端矩阵有逆. 消去过渡层 $\boldsymbol{e}^{(k+\frac{1}{2})}$, 得递推式:

$$\begin{aligned}&\boldsymbol{e}^{(k+1)}=\boldsymbol{T}_k\boldsymbol{e}^{(k)}=\cdots=\boldsymbol{T}_k\boldsymbol{T}_{k-1}\cdots\boldsymbol{T}_0\boldsymbol{e}^{(0)},\\ &\quad\boldsymbol{T}_k=(\boldsymbol{I}+\tau_k\boldsymbol{L}_2)^{-1}(\boldsymbol{I}-\tau_k\boldsymbol{L}_1)(\boldsymbol{I}+\tau_k\boldsymbol{L}_1)^{-1}(\boldsymbol{I}-\tau_k\boldsymbol{L}_2),\quad k=0,1,\cdots.\end{aligned}$$

由 §6 例 6.1, 可知矩阵 $\boldsymbol{L}_1$, $\boldsymbol{L}_2$ 有相同的特征向量 $\boldsymbol{v}^{(l,m)}=\{v_{ij}^{(l,m)}\}$:

$$v_{ij}^{(l,m)}=\sin li\pi h\quad\sin mj\pi h,\quad l,m,i,j=1,2,\cdots,N-1.$$

相应的特征值依次为

$$\frac{4}{h^2}\sin^2\frac{l\pi h}{2},\quad \frac{4}{h^2}\sin^2\frac{m\pi h}{2}.$$

于是 $\boldsymbol{L}_1$ 和 $\boldsymbol{L}_2$ 的乘积可换序, 从而 $\boldsymbol{T}_k$ 是对称矩阵, 其特征值

$$\begin{aligned}\lambda_{lm}^{(k)}&=\frac{\left(1-4\tau_kh^{-2}\sin^2\frac{1}{2}l\pi h\right)\left(1-4\tau_kh^{-2}\sin^2\frac{1}{2}m\pi h\right)}{\left(1+4\tau_kh^{-2}\sin^2\frac{1}{2}l\pi h\right)\left(1+4\tau_kh^{-2}\sin^2\frac{1}{2}m\pi h\right)},\\ &\quad l,m=1,2,\cdots,N-1.\end{aligned} \tag{2.7.5}$$

这样,

$$\|\boldsymbol{e}^{(k+1)}\|\leqslant\|\boldsymbol{T}_k\boldsymbol{T}_{k-1}\cdots\boldsymbol{T}_0\|\ \|\boldsymbol{e}^{(0)}\|=\max_{l,m}\prod_{j=0}^{k}|\lambda_{lm}^{(j)}|\,\|\boldsymbol{e}^{(0)}\|. \tag{2.7.6}$$

我们的目的是选 $\tau_j(j=0,1,\cdots,k)$ 使右端取极小. 精确求 τ_j 是不可能的, 只能近似求解.

(i) 单参数情形. 此时 $\tau_k=\tau$. 令 $\mu=4\tau/h^2$, 选 μ 使

$$\max_{a\leqslant x\leqslant b}\left|\frac{1-\mu x}{1+\mu x}\right|=\min,$$

其中

$$a=\sin^2\frac{\pi h}{2},\quad b=\cos^2\frac{\pi h}{2}.$$

因 $\mu>0$, 函数 $(1-\mu x)(1+\mu x)^{-1}$ 关于 x 递减. 又对固定 $x>0$, 函数 $(1-\mu x)(1+\mu x)^{-1}$ 关于 μ 递减. 故应选 μ 使 $(1-\mu x)(1+\mu x)^{-1}$ 在区间 $[a,b]$ 的端点绝对值相等符号相反, 即

$$\max_{a\leqslant x\leqslant b}\left|\frac{1-\mu x}{1+\mu x}\right|=\frac{1-\mu a}{1+\mu a}=-\frac{1-\mu b}{1+\mu b}.$$

由此解得

$$\mu_{opt}=(ab)^{-\frac{1}{2}}, \tag{2.7.7}$$

$$\tau_{opt}=\frac{1}{2}h^2(\sin\pi h)^{-1}. \tag{2.7.8}$$

此时

$$\begin{aligned}\max_{l,m}|\lambda_{lm}|&=\left(\frac{1-\mu_{opt}a}{1+\mu_{opt}a}\right)^2=\left(\frac{1-\operatorname{tg}\frac{1}{2}\pi h}{1+\operatorname{tg}\frac{1}{2}\pi h}\right)^2\\&=\frac{1-\sin\pi h}{1+\sin\pi h}\sim 1-2\pi h,\quad h\to 0,\\R(\boldsymbol{T})&=-\ln\frac{1-\sin\pi h}{1+\sin\pi h}\sim 2\pi h,\quad h\to 0.\end{aligned}$$

敛速和 SOR 方法相同 (见 (2.6.34)).

(ii) p 个参数情形. 选参数 $\mu_k=4\tau_k h^{-2}$, $k=0,1,\cdots,p-1$, 使

$$\max_{a\leqslant x\leqslant b}\left|\prod_{k=0}^{p-1}\frac{1-\mu_k x}{1+\mu_k x}\right|=\min.$$

对区间 $[a,b]$ 作分割:

$$a=\xi_0<\xi_1<\cdots<\xi_{p-1}<\xi_p=b.$$

节点 ξ_k 待定. 显然, 对 $\forall k=0,1,\cdots,p-1$

$$\max_{a\leqslant x\leqslant b}\left|\prod_{k=0}^{p-1}\frac{1-\mu_k x}{1+\mu_k x}\right|\leqslant\max_{\xi_{k-1}\leqslant x\leqslant\xi_k}\left|\frac{1-\mu_{k-1}x}{1+\mu_{k-1}x}\right|. \tag{2.7.9}$$

按 (i) 求 μ_{k-1} 使

$$\max_{\xi_{k-1}\leqslant x\leqslant\xi_k}\left|\frac{1-\mu_k x}{1+\mu_k x}\right|=\min.$$

由 (2.7.7), $\mu_{k-1}=(\xi_{k-1}\xi_k)^{-\frac{1}{2}}$, 相应的

$$\max_{\xi_{k-1}\leqslant x\leqslant\xi_k}\left|\frac{1-\mu_{k-1}x}{1+\mu_{k-1}x}\right|=\frac{1-(\xi_{k-1}/\xi_k)^{\frac{1}{2}}}{1+(\xi_{k-1}/\xi_k)^{\frac{1}{2}}}.$$

今选节点 ξ_k $(k=0,1,\cdots,p)$, 使

$$\frac{1-(\xi_{k-1}/\xi_k)^{\frac{1}{2}}}{1+(\xi_{k-1}/\xi_k)^{\frac{1}{2}}}=\rho, \quad k=1,2,\cdots,p,$$

ρ 为一与 k 无关的待定系数, 由此解出

$$\xi_k=\left(\frac{1+\rho}{1-\rho}\right)^2\xi_{k-1}=\cdots=\left(\frac{1+\rho}{1-\rho}\right)^{2k}a, \quad k=1,2,\cdots,p,$$

特别 $b=\xi_p=(1+\rho)^{2p}(1-\rho)^{-2p}a$. 而 $a=\sin^2\dfrac{\pi h}{2}$, $b=\cos^2\dfrac{\pi h}{2}$, 所以

$$\begin{aligned}(2.7.10)\qquad p&=\ln\left(\frac{b}{a}\right)^{\frac{1}{2}}\Big/\ln\frac{1+\rho}{1-\rho}=-\ln\left(\mathrm{tg}\frac{\pi h}{2}\right)\Big/\ln\frac{1+\rho}{1-\rho}\\&\approx\left(\ln\frac{1+\rho}{1-\rho}\right)^{-1}\ln\frac{2}{\pi h}, \quad h\to 0.\end{aligned}$$

取定 p 和参数

$$\begin{aligned}(2.7.11)\qquad \mu_{k-1}&=(\xi_{k-1}\xi_k)^{-\frac{1}{2}}=a^{-1}\left(\frac{1-\rho}{1+\rho}\right)^{2k-1},\\ \tau_{k-1}&=\frac{h^2}{4}\mu_{k-1}=\frac{h^2}{4a}\left(\frac{1-\rho}{1+\rho}\right)^{2k-1}, \quad k=1,2,\cdots,p,\end{aligned}$$

循环用参数 $\tau_0,\tau_1,\cdots,\tau_{p-1}$ 作 PR 迭代. 由 (2.7.5), (2.7.6) 和 (2.7.9), 每循环一次, 初始误差缩减 ρ^2 倍, 因此平均敛速

$$\begin{aligned}(2.7.12)\qquad R_{PR}&=(-\ln\rho^2)\big/p=-2\ln\frac{1}{\rho}\ln\frac{1+\rho}{1-\rho}\Big/\ln\mathrm{tg}\frac{\pi h}{2}\\&\approx 2\ln\frac{1}{\rho}\ln\frac{1+\rho}{1-\rho}\Big/\ln\frac{2}{\pi h}=O\left(\ln^{-1}\frac{1}{h}\right), \quad h\to 0.\end{aligned}$$

ρ 应该选成使上式右端的因子 $\ln\dfrac{1}{\rho}\ln(1+\rho)(1-\rho)^{-1}$ 取最大, 为此 ρ 应满足 $\rho^{-1}=(1+\rho)/(1-\rho)$, 从而 $\rho=\sqrt{2}-1$, 故最大敛速

$$(2.7.13)\qquad R_{\mathrm{PR}}=-\frac{2\ln^2(\sqrt{2}-1)}{\ln\mathrm{tg}\dfrac{\pi h}{2}}>\frac{1.554}{\ln\dfrac{2}{\pi h}}=O\left(\log^{-1}\frac{1}{h}\right), \quad h\to 0.$$

比 SOR 的敛速 $(R_{\mathrm{SOR}}\sim 2\pi h)$ 有实质性改进.

在以上的推导中, 用到矩阵 $\boldsymbol{L}_i$ 之间乘积可换序的条件, 这当然是苛刻的. 实际遇到的问题不满足交换性条件 (例如区域非乘积型), 只好选取近似参数. 关于 $\boldsymbol{L}_1$, $\boldsymbol{L}_2$ 不可交换时的交替方向迭代法, 在 [32] 的第七章及其评注中有讨论.

习 题

1. 在交替方向迭代 (2.7.3) 中, 假定矩阵 $\boldsymbol{L}_1$, $\boldsymbol{L}_2$ 对称正定但乘积不必可交换, 试就单参数 $\tau_k = \tau > 0$ 证明方法收敛并确定最佳参数.

2. 设有方程组

$$\begin{aligned} &-a_i u_{i-1} + b_i u_i - c_i u_{i+1} = d_i, \quad i = 1, 2, \cdots, N-1, \\ &u_0 = u_N = 0, \end{aligned}$$

$b_i \neq 0$ $(i = 1, 2, \cdots, N-1)$, 试导出下列求解公式:

$$\begin{aligned} \beta_k &= \frac{c_k}{b_k - a_k \beta_{k-1}}, \beta_0 = 0, \\ v_k &= \frac{d_k + a_k v_{k-1}}{b_k - a_k \beta_{k-1}}, v_0 = 0 \quad (k = 1, 2, \cdots, N-1), \\ u_j &= \beta_j u_{j+1} + v_j, u_N = 0 \quad (j = N-1, N-2, \cdots, 1). \end{aligned}$$

3. (实习题) 用 PR 法求解 §6 的实习题 3

$$\begin{aligned} &\Delta u = 2\pi^2(\sin \pi x \cos \pi y + \cos \pi x \sin \pi y) e^{\pi(x+y)}, (x, y) \in R = (0, 1) \times (0, 1), \\ &\quad u = 0, (x, y) \in \partial R. \end{aligned}$$

其精确解为 $u(x, y) = e^{\pi(x+y)} \sin \pi x \sin \pi y$. 取步长 $h = k = 1/64, 1/128$, 作五点差分格式. 用单参数和双参数 PR 法解差分方程, 近似到小数点后四位. 与 SOR 法比较精度和迭代步数.

§8 预处理共轭梯度法

共轭梯度法是一种无需选松弛因子的迭代法. 如果计算是精确的, 则经有限步 (不超过方程组的阶数) 就可求到精确解, 所以也是直接法. 但计算中的舍入误差是不可避免的, 迭代收敛可能很慢, 故这一方法有一段时间未受重视. 到 20 世纪 70 年代, 由于采用预处理技术, 共轭梯度法已成为求解稀疏矩阵方程的有效算法.

8.1 共轭梯度法

经典的共轭梯度法已在 "数值代数" 教程中讲过, 这里只列出计算过程和结果 (参看 [4]). 设有对称正定系数矩阵的 N 阶方程组:

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{f}. \tag{2.8.1}$$

用共轭梯度法求解 (2.8.1) 的步骤如下: 取初值 $\boldsymbol{u}^{(0)}$, 计算

$$
\begin{aligned}
&\boldsymbol{p}^{(0)}=\boldsymbol{r}^{(0)}=\boldsymbol{f}-\boldsymbol{A}\boldsymbol{u}^{(0)},\\
&\alpha_m=\frac{(\boldsymbol{r}^{(m)},\boldsymbol{r}^{(m)})}{(\boldsymbol{A}\boldsymbol{p}^{(m)},\boldsymbol{p}^{(m)})},\\
&\boldsymbol{u}^{(m+1)}=\boldsymbol{u}^{(m)}+\alpha_m\boldsymbol{p}^{(m)},\\
&\boldsymbol{r}^{(m+1)}=\boldsymbol{r}^{(m)}-\alpha_m\boldsymbol{A}\boldsymbol{p}^{(m)},\\
&\beta_m=\frac{(\boldsymbol{r}^{(m+1)},\boldsymbol{r}^{(m+1)})}{(\boldsymbol{r}^{(m)},\boldsymbol{r}^{(m)})},\\
&\boldsymbol{p}^{(m+1)}=\boldsymbol{r}^{(m+1)}+\beta_m\boldsymbol{p}^{(m)},
\end{aligned}
\tag{2.8.2}
$$

$m=1,2,\cdots,M$. 若计算是精确的, 则有 $M<N$ 使 $\boldsymbol{u}^{(M)}=\boldsymbol{u}$, $\boldsymbol{u}$ 是 (2.8.1) 的解. 按 (2.8.2) 决定的两组向量 $\boldsymbol{r}^{(i)}$, $\boldsymbol{p}^{(i)}$ 有下列性质:

$$(\boldsymbol{r}^{(i)},\boldsymbol{r}^{(j)})=0,\quad i\neq j,\tag{2.8.3}$$

$$(\boldsymbol{p}^{(i)},\boldsymbol{A}\boldsymbol{p}^{(j)})=0,\quad i\neq j,\tag{2.8.4}$$

$$(\boldsymbol{p}^{(i)},\boldsymbol{r}^{(j)})=0,\quad i<j,\tag{2.8.5}$$

$$(\boldsymbol{p}^{(i)},\boldsymbol{r}^{(j)})=(\boldsymbol{r}^{(i)},\boldsymbol{r}^{(i)}),\quad j\leqslant i,\tag{2.8.6}$$

$$(\boldsymbol{r}^{(i)},\boldsymbol{A}\boldsymbol{p}^{(i)})=(\boldsymbol{p}^{(i)},\boldsymbol{A}\boldsymbol{p}^{(i)}),\tag{2.8.7}$$

$$(\boldsymbol{r}^{(i)},\boldsymbol{A}\boldsymbol{p}^{(j)})=0,\quad i\neq j,j+1.\tag{2.8.8}$$

设 $\boldsymbol{A}$ 的特征值为

$$0<\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_N,$$

则 $\boldsymbol{A}$ 的条件数

$$\mathrm{Cond}(\boldsymbol{A})=\lambda_N/\lambda_1.$$

第 m 次迭代 $\boldsymbol{u}^{(m)}$ 有误差估计:

$$
\begin{aligned}
\|\boldsymbol{u}^{(m)}-\boldsymbol{u}\|_{\boldsymbol{A}}&\leqslant 2\left(\frac{\sqrt{\lambda_N}-\sqrt{\lambda_1}}{\sqrt{\lambda_N}+\sqrt{\lambda_1}}\right)^m\|\boldsymbol{u}-\boldsymbol{u}^{(0)}\|_{\boldsymbol{A}}\\
&=2\left(\frac{\sqrt{\mathrm{Cond}(\boldsymbol{A})}-1}{\sqrt{\mathrm{Cond}(\boldsymbol{A})}+1}\right)^m\|\boldsymbol{u}-\boldsymbol{u}^{(0)}\|_{\boldsymbol{A}},
\end{aligned}
\tag{2.8.9}
$$

此处 $\|\boldsymbol{x}\|_{\boldsymbol{A}}=\sqrt{(\boldsymbol{A}\boldsymbol{x},\boldsymbol{x})}$. 为使误差缩减成初始误差的 $\varepsilon\ (>0)$ 倍, 应取 m 使

$$m\geqslant\frac{1}{2}\sqrt{\mathrm{Cond}(\boldsymbol{A})}\ln\frac{2}{\varepsilon}.$$

将共轭梯度法用于五点差分格式 (2.6.9), 则

$$\mathrm{Cond}(\boldsymbol{A}) \sim \frac{4}{\pi^2 h^2}, \sqrt{\mathrm{Cond}(\boldsymbol{A})} \sim \frac{2}{\pi h},$$
$$m \geqslant \frac{1}{\pi h} \ln \frac{2}{\varepsilon}.$$

足见敛速与 SOR 法同阶.

8.2 预处理共轭梯度法

为了提高敛速, 克服数值不稳定性, 需对矩阵 $\boldsymbol{A}$ 作预处理, 减小问题的条件数, 然后用共轭梯度法求解, 这就是**预处理共轭梯度法**.

取某一对称正定矩阵 $\boldsymbol{B}$, 将 (2.8.1) 化为

$$\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{u} = \boldsymbol{B}^{-1}\boldsymbol{f}. \tag{2.8.10}$$

令

$$\widehat{\boldsymbol{A}} = \boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{B}^{-\frac{1}{2}}, \quad \widehat{\boldsymbol{u}} = \boldsymbol{B}^{\frac{1}{2}}\boldsymbol{u}, \quad \widehat{\boldsymbol{f}} = \boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{f}, \tag{2.8.11}$$

将 (2.8.10) 写成

$$\widehat{\boldsymbol{A}}\widehat{\boldsymbol{u}} = \widehat{\boldsymbol{f}}. \tag{2.8.12}$$

此时 $\widehat{\boldsymbol{A}}$ 仍为对称正定矩阵. 若经 (2.8.11) 变换后的 $\widehat{\boldsymbol{A}}$ 满足

$$\mathrm{Cond}(\widehat{\boldsymbol{A}}) << \mathrm{Cond}(\boldsymbol{A}),$$

则称如此的 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的**预处理器**(Preconditioner).

将共轭梯度法用于处理后的方程 (2.8.12), 得到一套和 (2.8.2) 平行的计算公式, 只需将所有向量和矩阵 $\boldsymbol{A}$ 都戴帽 $\wedge$. 然后用两套变量间的关系

$$\widehat{\boldsymbol{u}}^{(m)} = \boldsymbol{B}^{\frac{1}{2}}\boldsymbol{u}^{(m)}, \ \widehat{\boldsymbol{r}}^{(m)} = \boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{r}^{(m)}, \ \widehat{\boldsymbol{p}}^{(m)} = \boldsymbol{B}^{\frac{1}{2}}\boldsymbol{p}^{(m)},$$
$$\widehat{\boldsymbol{A}} = \boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{B}^{-\frac{1}{2}}, \widehat{\boldsymbol{f}} = \boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{f},$$

并引进变量

$$\boldsymbol{s}^{(m)} = \boldsymbol{B}^{-1}\boldsymbol{r}^{(m)} \text{ 或 } \boldsymbol{B}\boldsymbol{s}^{(m)} = \boldsymbol{r}^{(m)}, \tag{2.8.13}$$

则得到关于 $\boldsymbol{u}^{(m)}$ 的计算公式: 取初值 $\boldsymbol{u}^{(0)}$, 计算

$$(2.8.14)_1 \qquad \boldsymbol{p}^{(0)} = \boldsymbol{r}^{(0)} = \boldsymbol{f} - \boldsymbol{A}\boldsymbol{u}^{(0)},$$

$$(2.8.14)_2 \qquad \boldsymbol{s}^{(0)} = \boldsymbol{B}^{-1}\boldsymbol{r}^{(0)} \quad (\boldsymbol{B}\boldsymbol{s}^{(0)} = \boldsymbol{r}^{(0)}),$$

$$(2.8.14)_3 \qquad \alpha_m = \frac{(\boldsymbol{r}^{(m)}, \boldsymbol{s}^{(m)})}{(\boldsymbol{A}\boldsymbol{p}^{(m)}, \boldsymbol{p}^{(m)})},$$

$$(2.8.14)_4 \qquad \boldsymbol{u}^{(m+1)} = \boldsymbol{u}^{(m)} + \alpha_m \boldsymbol{p}^{(m)},$$

$$(2.8.14)_5 \qquad \boldsymbol{r}^{(m+1)} = \boldsymbol{r}^{(m)} - \alpha_m \boldsymbol{A}\boldsymbol{p}^{(m)},$$

$$(2.8.14)_6 \qquad \boldsymbol{s}^{(m+1)} = \boldsymbol{B}^{-1}\boldsymbol{r}^{(m+1)} \quad (\boldsymbol{B}\boldsymbol{s}^{(m+1)} = \boldsymbol{r}^{(m+1)}),$$

$$(2.8.14)_7 \qquad \beta_m = \frac{(\boldsymbol{r}^{(m+1)}, \boldsymbol{s}^{(m+1)})}{(\boldsymbol{r}^{(m)}, \boldsymbol{s}^{(m)})},$$

$$(2.8.14)_8 \qquad \boldsymbol{p}^{(m+1)} = \boldsymbol{s}^{(m+1)} + \beta_m \boldsymbol{p}^{(m)}.$$

与 (2.8.2) 比较, 预处理共轭梯度法多了一步 $(2.8.14)_6$, 若预处理器 $\boldsymbol{B}$ 使 $\boldsymbol{B}\boldsymbol{s}^{(m)} = \boldsymbol{r}^{(m)}$ 易于求解, 则实现 $(2.8.14)_6$ 的工作量不大.

注意

$$\|\widehat{\boldsymbol{u}}^{(m)} - \widehat{\boldsymbol{u}}\|_{\widehat{\boldsymbol{A}}} = \|\boldsymbol{u}^{(m)} - \boldsymbol{u}\|_{\boldsymbol{A}},$$

若 $\mathrm{Cond}(\widehat{\boldsymbol{A}}) << \mathrm{Cond}(\boldsymbol{A})$, 则预处理共轭梯度法的敛速将显著提高. 问题是如何构造预处理矩阵 $\boldsymbol{B}$, 既使 $\mathrm{Cond}(\widehat{\boldsymbol{A}})$ 明显减小, 又使 $(2.8.14)_6$ 容易实现. 迄今已发展了一系列有效的预处理技术 (参看 [11]).

在理想情形, 设 $\boldsymbol{A}$ 能完全分解为 $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$, 其中 $\boldsymbol{L}$ 和 $\boldsymbol{U}$ 是下三角和上三角矩阵, 则可取 $\boldsymbol{B} = \boldsymbol{L}\boldsymbol{U}$, 此时 $\boldsymbol{B}^{-1}\boldsymbol{A}$ 是单位矩阵, 条件数达到极小. 在一般情形, 可将 $\boldsymbol{A}$ 作不完全分解:

$$(2.8.15) \qquad \boldsymbol{A} = \boldsymbol{B} + \boldsymbol{R},$$

其中 $\boldsymbol{B} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}}$, $\boldsymbol{D}$ 是对角阵, $\boldsymbol{L}$ 是下三角阵, $\boldsymbol{L}^{\mathrm{T}}$ 是 $\boldsymbol{L}$ 的转置; $\boldsymbol{R}$ 是剩余矩阵 (假定在某种意义下较小), 如此的 $\boldsymbol{B}$ 可取作预处理阵. 实现不完全分解 (2.8.15) 的方案很多, 现以五点差分格式为例介绍由 Meijerink 和 Vorst 提出的不完全因子分解法 ([11]).

将差分方程 (2.6.27) 按照由左到右, 由下到上的次序排列, 并作统一编号 $n = 1, 2, \cdots$. 网格函数用向量 $\boldsymbol{u} = (u_1, u_2, \cdots)^{\mathrm{T}}$ 表示, $\boldsymbol{A}$ 是相应的系数矩阵. 若

更设 $\boldsymbol{A}$ 对称, 将 (2.6.27) 写成:

$$(\boldsymbol{A}\boldsymbol{u})_n = a_{n-N,n}u_{n-N} + a_{n-1,n}u_{n-1} + a_{n,n}u_n + a_{n,n+1}u_{n+1} + a_{n,n+N}u_{n+N}, \tag{2.8.16}$$

$$n = 1, 2, \cdots, M \quad (M = (N-1)^2);$$

$$a_{jn} = 0 \quad \text{当 } j \leqslant 0, \quad a_{nj} = 0 \quad \text{当 } j > M.$$

矩阵 $\boldsymbol{A}$ 的非零元素分布在五个斜对角线上, 如图 8.1(a); 差分格式的网点系数分布如图 8.1(b).

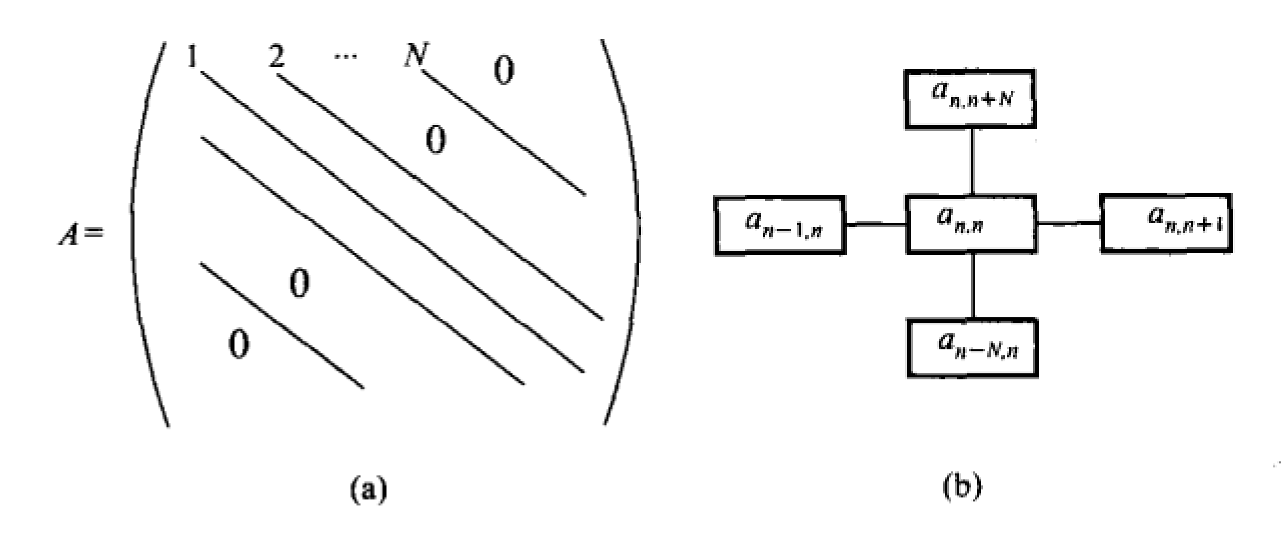


图 8.1

一个最简单的不完全因子分解法是取 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的三对角阵:

$$\boldsymbol{B} = \begin{bmatrix} a_{11} & a_{12} & & 0 \\ a_{12} & a_{22} & \ddots & \\ & \ddots & \ddots & a_{M-1,M} \\ 0 & & a_{M-1,M} & a_{M,M} \end{bmatrix}$$

用消元法对 $\boldsymbol{B}$ 作 $\boldsymbol{L}, \boldsymbol{U}$ 分解.

更常用的是按 Meijerink-Vorst 的方法对 $\boldsymbol{A}$ 作不完全因子分解. 取 $\boldsymbol{D} = \operatorname{diag}(d_1, d_2, \cdots, d_M)$, $\boldsymbol{L}$ 为单位下三角阵:

$$(\boldsymbol{L}\boldsymbol{u})_n = u_n + l_{n-1,n}u_{n-1} + l_{n-N,n}u_{n-N}, \tag{2.8.17}$$

$l_{jn} = 0$ 当 $j \leqslant 0$. $\boldsymbol{L}$ 的结构如图 8.2, 其中非零元素的分布和 $\boldsymbol{A}$ 的下三角阵相同, 但数值待定. 由 (2.8.17) 知

$$(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{u})_n = u_n + l_{n,n+1}u_{n+1} + l_{n,n+N}u_{n+N}, \tag{2.8.18}$$

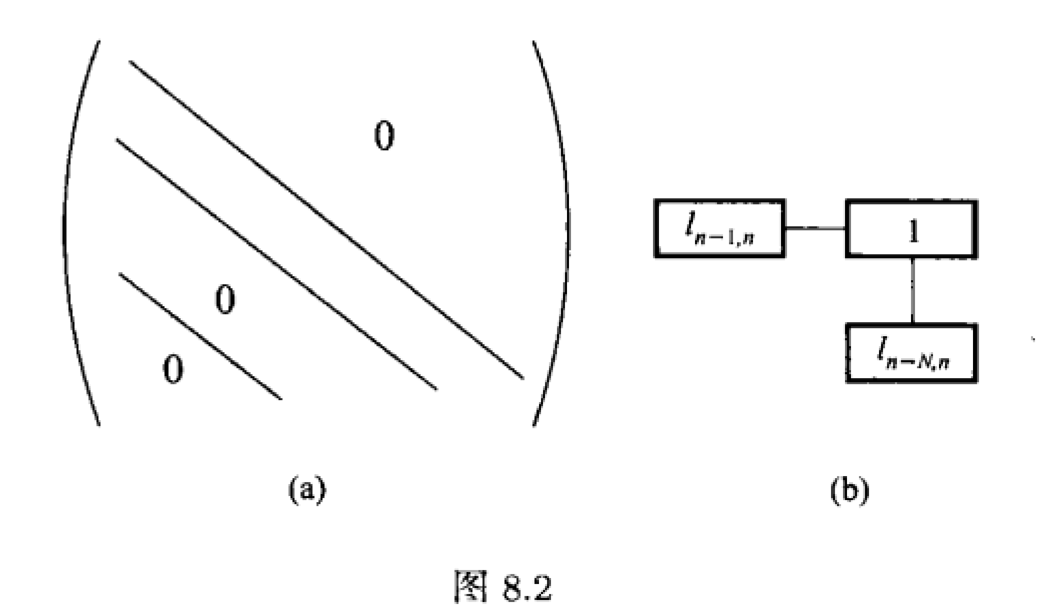


图 8.2

于是

$$
\begin{aligned}
&(\boldsymbol{DL^T u})_n = d_n(u_n + l_{n,n+1}u_{n+1} + l_{n,n+N}u_{n+N}),\\
(2.8.19)\quad &(\boldsymbol{L}(\boldsymbol{DL^T u}))_n = (\boldsymbol{DL}^{\mathrm{T}}\boldsymbol{u})_n + l_{n-1,n}(\boldsymbol{DL}^{\mathrm{T}}\boldsymbol{u})_{n-1} + l_{n-N,n}(\boldsymbol{DL}^{\mathrm{T}}\boldsymbol{u})_{n-N}\\
&= (d_n + d_{n-1}l_{n-1,n}^2 + d_{n-N}l_{n-N,n}^2)u_n +\\
&\quad d_n(l_{n,n+1}u_{n+1} + l_{n,n+N}u_{n+N}) +\\
&\quad d_{n-1}l_{n-1,n}(u_{n-1} + l_{n-1,n+N-1}u_{n+N-1}) +\\
&\quad d_{n-N}l_{n-N,n}(u_{n-N} + l_{n-N,n-N+1}u_{n-N+1})
\end{aligned}
$$

矩阵 $\boldsymbol{LDL}^{\mathrm{T}}$ 每行可能有七个非零元素, 如图 8.3(a), 相应的差分算子的结构如图 8.3(b), 多了右下角和左上角两个相关节点. 令 (2.8.19) 中 u_n, u_{n+1}, u_{n+N} 的

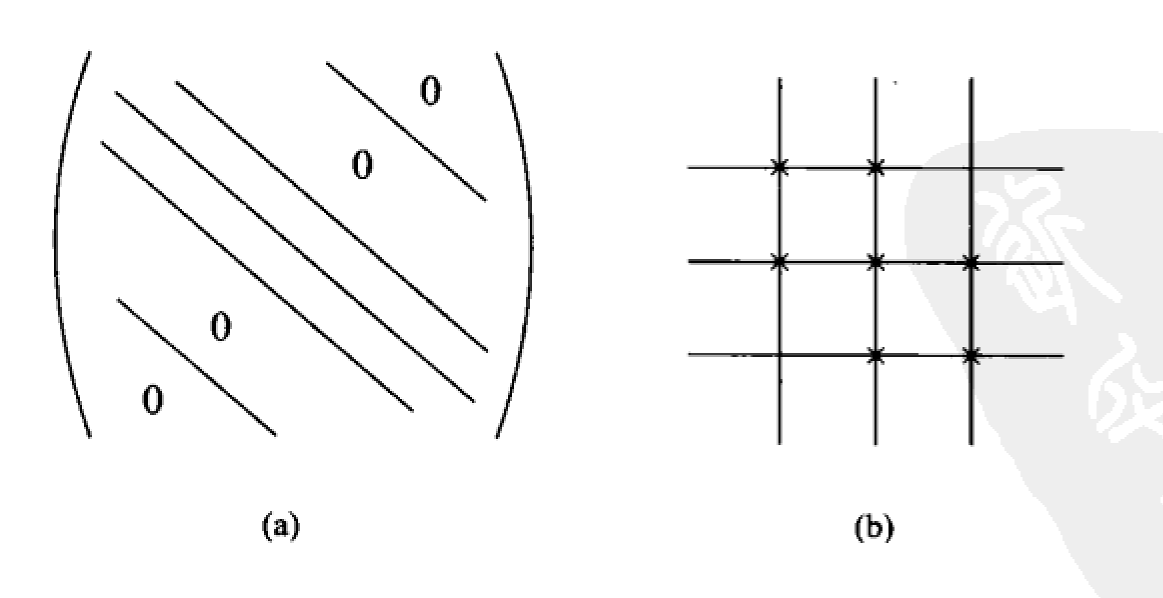


图 8.3

系数与 (2.8.16) 的相应系数相等 (由对称性可推出 u_{n-1}, u_{n-N} 的相应系数也相

等), 则得

(2.8.20) $d_n = a_{nn} - d_{n-1}l_{n-1,n}^2 - d_{n-N}l_{n-N,n}^2, \quad n = 1, 2, \cdots, M,$

(2.8.21) $l_{n,n+1} = a_{n,n+1}/d_n, \quad n = 1, 2, \cdots, M-1,$

(2.8.22) $l_{n,n+N} = a_{n,n+N}/d_n, \quad n = 1, 2, \cdots, M-N,$

其中 $d_j = l_{jn} = 0$, 当 $j \leqslant 0$; $a_{nj} = l_{nj} = 0$, 当 $j > M$.

如上确定的 $\boldsymbol{B} = \boldsymbol{LDL}^{\mathrm{T}}$ 就是所需的预处理阵. 显然矩阵 $\boldsymbol{B}$ 容易计算. 大量计算表明, $\boldsymbol{B}^{-1}\boldsymbol{A}$ 的条件数比 $\boldsymbol{A}$ 的条件数有很大改善. 需指出的是, 如此做的分解 $\boldsymbol{A} = \boldsymbol{B} + \boldsymbol{R}$ 对应一种 "强隐式迭代法":

$$\boldsymbol{Bu}^{(m+1)} = -\boldsymbol{Ru}^{(m)} + \boldsymbol{f}$$

或

$$\boldsymbol{u}^{(m+1)} = -\boldsymbol{B}^{-1}\boldsymbol{Ru}^{(m)} + \boldsymbol{B}^{-1}\boldsymbol{f}.$$

实际上, 每一种迭代都对应 $\boldsymbol{A}$ 的一种不完全因子分解, 所以也提供一种预处理共轭梯度法. 近年发展的用对称超松弛法 (见 §6 注) 建立的预处理共轭梯度法就是一种有效的算法 (参看 [19]).

习　　题

(实习题) 求解边值问题 (§6 的习题 3):

$$\Delta u = 2\pi^2 e^{\pi(x+y)}(\sin \pi x \cos \pi y + \cos \pi x \sin \pi y), \quad (x, y) \in R = (0, 1) \times (0, 1),$$
$$u = 0, \quad (x, y) \in \partial R.$$

其精确解为 $u(x, y) = e^{\pi(x+y)} \sin \pi x \sin \pi y$. 取步长 $h = k = 1/64, 1/128$, 作五点差分格式. 用共轭梯度法和预处理共轭梯度法解差分方程, 近似到小数点后四位. 与 SOR 法和 PR 法比较精度和迭代步数.

§9 数 值 例 子

用差分法求解边值问题:

$$-\Delta u = \cos 3x \sin \pi y, \quad (x, y) \in G = (0, \pi) \times (0, 1),$$
$$u(x, 0) = u(x, 1) = 0, \quad 0 \leqslant x \leqslant 1,$$
$$u_x(0, y) = u_x(\pi, y) = 0, \quad 0 \leqslant y \leqslant 1$$

(精确解 $u = (9 + \pi^2)^{-1} \cos 3x \sin \pi y$).

以步长 $h_1 = \dfrac{\pi}{N}$, $h_2 = \dfrac{1}{N}$ 作矩形剖分, 网格节点为 $x_i = ih_1$, $y_j = jh_2$, $i, j = 0, 1, \cdots, N$. 差分方程为

$$-\left(\frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h_1^2} + \frac{u_{i,j+1} - 2u_{ij} + u_{i,j-1}}{{h_2}^2}\right) = \cos 3x_i \sin \pi y_j,$$
$$i, j = 1, 2, \cdots, N-1.$$

边值条件为

$$u_{i0} = u_{iN} = 0, i = 0, \cdots, N,$$
$$u_{0j} = u_{1j}, j = 1, \cdots, N-1,$$
$$u_{Nj} = u_{N-1,j}, j = 1, \cdots, N-1.$$

求解方案: 依次令 $N = 4, 8, 16, 32$, 取 6 位小数计算.

(一) 用消元法求解, 并就 $(x_i, y_j) = \left(\dfrac{i\pi}{4}, \dfrac{j}{4}\right), i, j = 1, 2, 3$ 处列出差分解与精确解,如表 9.1. 从表中看出, 差分解随网格步长减小趋于精确解.

表 9.1

h	(x_j, y_k)	x_1	x_2	x_3	误差阶
1/4	y_1	−0.045 481	0.000 000	0.045 481	1.0×10^{-1}
1/8		−0.031 658	0.000 000	0.031 658	1.0×10^{-2}
1/16		−0.028 146	0.000 000	0.028 146	1.0×10^{-2}
1/32		−0.027 121	0.000 000	0.027 121	1.0×10^{-3}
精确解		−0.026 498	0.000 000	0.026 498	
1/4	y_2	−0.064 319	0.000 000	0.064 319	1.0×10^{-1}
1/8		−0.044 771	0.000 000	0.044 771	1.0×10^{-2}
1/16		−0.039 805	0.000 000	0.039 805	1.0×10^{-2}
1/32		0.038 356	0.000 000	0.038 356	1.0×10^{-3}
精确解		−0.037 473	0.000 000	0.037 473	
1/4	y_3	−0.045 481	0.000 000	0.045 481	1.0×10^{-1}
1/8		−0.031 658	0.000 000	0.031 658	1.0×10^{-2}
1/16		−0.028 146	0.000 000	0.028 146	1.0×10^{-2}
1/32		−0.027 121	0.000 000	0.027 121	1.0×10^{-3}
精确解		−0.026 498	0.000 000	0.026 498	

其次, 就 $N = 32$, $y = 0.25, 0.5, 0.75$ 及 $i = 0, 2, 4, \cdots, 30, 32$ 画出差分解曲线. 如图 9.1 实线为精确解, 虚线为近似解.

(二) 取初值 $u_0 = 1$ 依次用 Jacobi 迭代法, Seidel 迭代法, SOR 迭代法 ($\varpi = 1.45$), PR 迭代法 (单参数) 和共轭梯度法解差分方程, 记录迭代步数. 结果

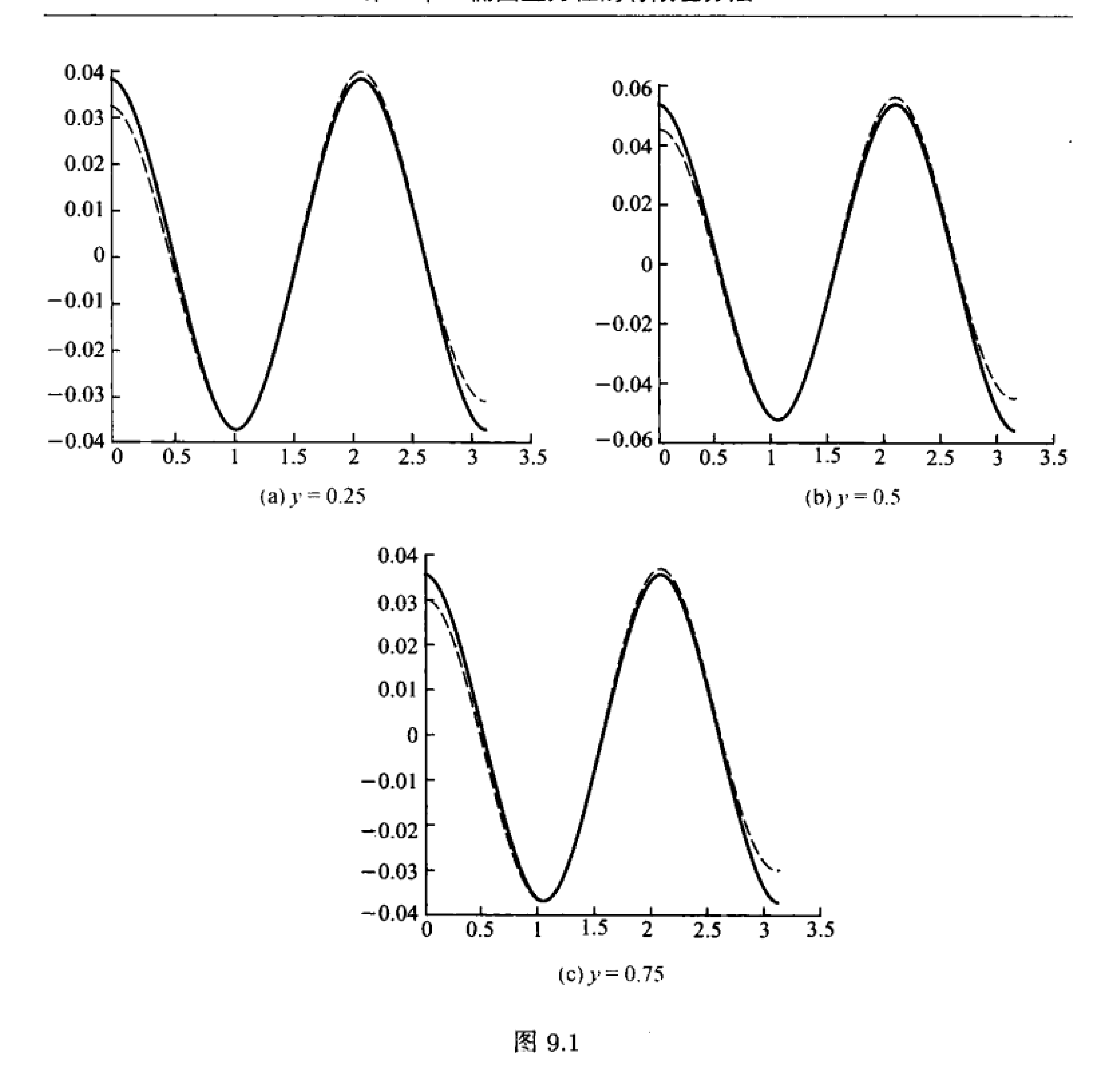


图 9.1

如表 9.2. 从中看出 Seidel 法的敛速确为 Jacobi 法的二倍, 而 SOR 法, PR 法和共轭梯度法较 Seidel 法快很多.

表 9.2

	h	Jacobi 法	Seidel 法	SOR 法	PR 法	共轭梯度法
迭代次数	1/4	42	23	21	17	5
k	1/8	160	86	28	34	9
	1/16	571	306	124	64	18
	1/32	1 974	1 065	446	120	42

第三章　抛物型方程的有限差分法

椭圆型方程描写的状态 (如温度、电位等) 不随时间 t 改变, 称为**驻定问题**. 现在讨论与时间 t 有关的**非驻定问题**. 驻定问题可看成是某一非驻定问题当 $t \to \infty$ 的渐近状态, 所以, 当我们用渐近方法 (例如迭代法) 求解驻定问题时, 只关心最终状态, 而不管中间过程. 相反, 非驻定问题的瞬时状态有物理意义, 需要我们求解. 在考虑偏微分方程的数值解法时, 注意到这两类问题的联系和区别是有益的. 本章和下一章分别讨论抛物方程和双曲方程的差分法.

§1　最简差分格式

作为模型, 考虑一维热传导方程:

$$\frac{\partial u}{\partial t} = a\frac{\partial^2 u}{\partial x^2} + f(x), \quad 0 < t \leqslant T, \tag{3.1.1}$$

其中 a 是正常数, $f(x)$ 是给定的连续函数. 按照初边值条件的不同给法, 可将 (3.1.1) 的定解问题分为两类:

第一, 初值问题 (也称 Cauchy 问题): 求具有一定阶偏微商的函数 $u(x,t)$, 满足方程 (3.1.1) 和初始条件:

$$u(x,0) = \phi(x), \quad -\infty < x < \infty. \tag{3.1.2}$$

第二, 初边值问题 (也称混合问题): 求具有一定阶偏微商的函数 $u(x,t)$, 满足方程 (3.1.1) 和初始条件:

$$u(x,0) = \phi(x), \quad 0 < x < l \tag{$3.1.3_1$}$$

及边值条件

$$u(0,t) = u(l,t) = 0, \quad 0 \leqslant t \leqslant T. \tag{$3.1.3_2$}$$

假定 $f(x)$ 和 $\phi(x)$ 在相应区域光滑, 并且在 $x = 0, l$ 相容, 则上述问题有唯一的光滑解.

现在考虑边值问题 (3.1.1), (3.1.3) 的差分逼近. 取空间步长 $h = l/J$ 和时间步长 $\tau = T/N$, 其中 J, N 都是自然数. 用两族平行直线 $x = x_j = jh (j = 0, 1, \cdots, J)$

和 $t=t_n=n\tau(n=0,1,\cdots,N)$ 将矩形域 $\overline{G}=\{0\leqslant x\leqslant l;0\leqslant t\leqslant T\}$ 分割成矩形网格, 网格节点为 (x_j,t_n). 以 G_h 表示网格内点集合, 即位于开矩形 G 的网点集合; $\overline{G}_h$ 表示所有位于闭矩形 $\overline{G}$ 的网点集合; $\Gamma_h=\overline{G}_h-G_h$ 是网格界点集合 (参看图 1.1).

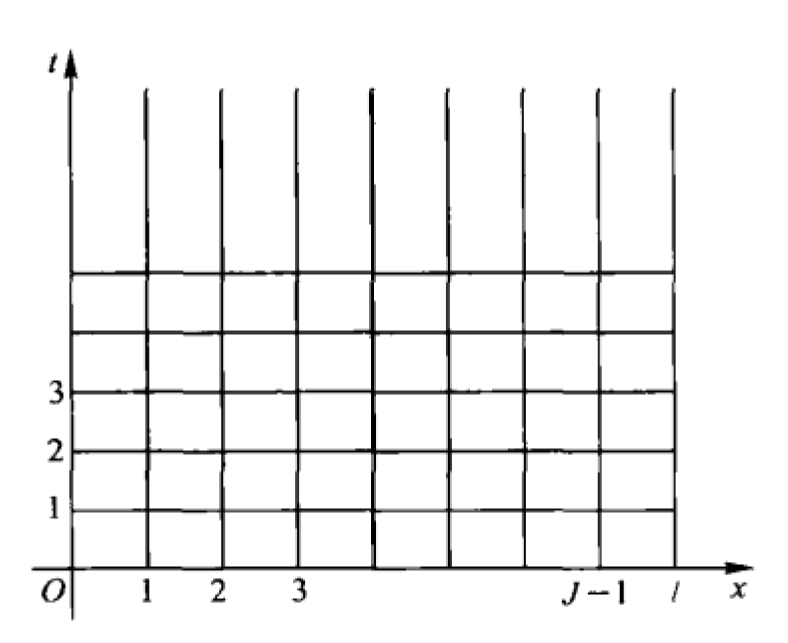


图 1.1

其次, 用 u_j^n 表示定义在网点 (x_j,t_n) 上的函数,$0\leqslant j\leqslant J,0\leqslant n\leqslant N$. 与第二章 §1类似, 用适当的差商代替方程 (3.1.1) 中相应的偏微商, 便得到以下几种最简差分格式.

(一) **向前差分格式**, 即

$$(3.1.4)_1\qquad \frac{u_j^{n+1}-u_j^n}{\tau}=a\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2}+f_j,$$

$$f_j=f(x_j),$$

$$(3.1.4)_2\qquad u_j^0=\phi_j=\phi(x_j),\quad u_0^n=u_J^n=0,$$

其中 $j=1,2,\cdots,J-1,n=0,1,\cdots,N-1$. 以 $r=a\tau/h^2$ 表示**网比**. 将 $(3.1.4)_1$ 改写成便于计算的形式, 使第 n 层值 (上标为 n) 在等式右边, 第 $n+1$ 层值在等式左边, 则得

$$(3.1.4)_1'\qquad u_j^{n+1}=ru_{j+1}^n+(1-2r)u_j^n+ru_{j-1}^n+\tau f_j.$$

取 $n=0$, 利用初值 $u_j^0=\varphi_j$ 和边值 $u_0^n=u_J^n=0$ 可由 $(3.1.4)_1'$ 算出第一层值 u_j^1. 于 $(3.1.4)_1'$ 取 $n=1$, 又可利用 u_j^1 和边值, 由 $(3.1.4)_1'$ 算出 u_j^2. 如此下去, 即可逐层求出所有 u_j^n, 并视 u_j^n 为精确解 $u(x_j,t_n)$ 的近似. 由于第 $(n+1)$ 层值通过第 n 层值明显表为 $(3.1.4)_1'$, 无需解线性代数方程组, 如此的差分格式称为**显格式**. 将 $(3.1.4)_1'$ 看成网点 (x_j,t_n) 处的差分方程, 它联系第 $(n+1)$ 层的点 (x_j,t_{n+1}) 和第 n 层的点 $(x_{j-1},t_n),(x_j,t_n)$ 及 (x_{j+1},t_n), 其分布如图 1.2 (a) 所示.

记

$$Lu = \frac{\partial u}{\partial t} - a\frac{\partial^2 u}{\partial x^2},$$

$$L_h^{(1)}u_j^n = \frac{u_j^{n+1} - u_j^n}{\tau} - a\frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{h^2}.$$

显然截断误差

$$(3.1.5) \qquad \begin{aligned} R_j^n(u) &= L_h^{(1)}u(x_j, t_n) - [Lu]_j^n \\ &= -\tau\left[\frac{1}{12r} - \frac{1}{2}\right]\left(\frac{\partial^2 \tilde{u}}{\partial t^2}\right)_j^n + O(\tau^2 + h^2) \\ &= O(\tau + h^2), \end{aligned}$$

其中 $\left(\frac{\partial^2 \tilde{u}}{\partial t^2}\right)_j^n$ 是 $\frac{\partial^2 u}{\partial t^2}$ 在矩形 $x_{j-1} < x < x_{j+1}, t_n < t < t_{n+1}$ 中的某点值.

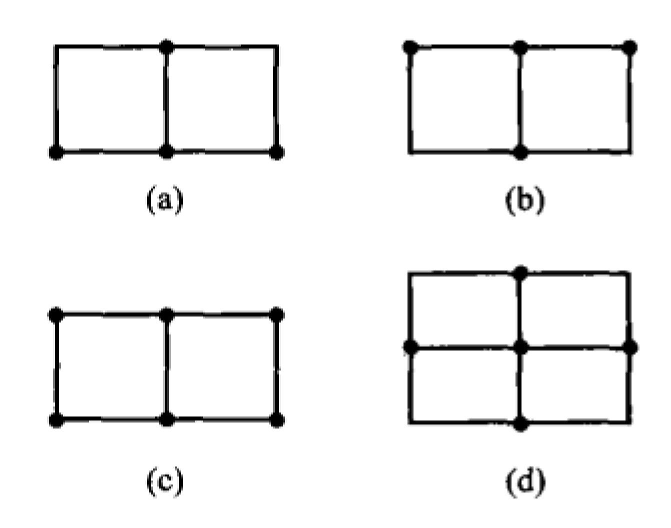


图 1.2

(二) **向后差分格式**, 即

$$(3.1.6)_1 \qquad \frac{u_j^{n+1} - u_j^n}{\tau} = a\frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{h^2} + f_j,$$

$$(3.1.6)_2 \qquad u_j^0 = \varphi_j = \varphi(x_j), \quad u_0^n = u_J^n = 0,$$

其中 $j = 1, 2, \cdots, J-1, n = 0, 1, \cdots, N-1$. 将 $(3.1.6)_1$ 改写为

$$(3.1.6)_1' \qquad -ru_{j+1}^{n+1} + (1+2r)u_j^{n+1} - ru_{j-1}^{n+1} = u_j^n + \tau f_j.$$

令 $n = 0, 1, 2, \cdots$, 则可利用 u_j^0 和边值确定 u_j^1, 利用 u_j^1 和边值确定 u_j^2, 等等. 现在第 $(n+1)$ 层的值不能用第 n 层值明显表示, 而是由线性代数方程组 $(3.1.6)_1'$ 确定, 如此的差分格式称为**隐格式**. 我们指出, $(3.1.6)_1'$ 左端系数矩阵严格对角占优, 方程总是可解的. 将 $(3.1.6)_1$ 看成是网点 (x_j, t_{n+1}) 处的差分方程, 它所联系的网点分布如图 1.2 (b) 所示.

令

$$L_h^{(2)}u_j^n = \frac{u_j^{n+1}-u_j^n}{\tau} - a\frac{u_{j+1}^{n+1}-2u_j^{n+1}+u_{j-1}^{n+1}}{h^2},$$

则截断误差

$$\begin{aligned}R_j^n(u) &= L_h^{(2)}u(x_j,t_n)-[Lu]_j^n \\ &= -\tau\left[\frac{1}{12r}+\frac{1}{2}\right]\left(\frac{\partial^2\tilde{u}}{\partial t^2}\right)_j^n + O(\tau^2+h^2) \\ &= O(\tau+h^2),\end{aligned} \tag{3.1.7}$$

其中 $\left(\frac{\partial^2\tilde{u}}{\partial t^2}\right)_j^n$ 是 $\frac{\partial^2 u}{\partial t^2}$ 在矩形 $x_{j-1}<x<x_{j+1}, t_n<t<t_{n+1}$ 中的某点值.

(二) **六点对称格式 (Crank–Nicolson 格式)**. 将向前差分格式和向后差分格式做算术平均, 即得六点对称格式:

$$\frac{u_j^{n+1}-u_j^n}{\tau} = \frac{a}{2}\left[\frac{u_{j+1}^{n+1}-2u_j^{n+1}+u_{j-1}^{n+1}}{h^2}+\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2}\right]+f_j, \tag{3.1.8$_1$}$$

$$u_j^0=\varphi_j=\varphi(x_j),\quad u_0^n=u_J^n=0. \tag{3.1.8$_2$}$$

将 $(3.1.8)_1$ 改写为

$$-\frac{r}{2}u_{j+1}^{n+1}+(1+r)u_j^{n+1}-\frac{r}{2}u_{j-1}^{n+1} = \frac{r}{2}u_{j+1}^n+(1-r)u_j^n+\frac{r}{2}u_{j-1}^n+\tau f_j, \tag{3.1.8$_1'$}$$

利用 u_j^0 和边值便可逐层求到 u_j^n. 六点对称格式是隐格式, 由第 n 层计算第 $(n+1)$ 层时, 需解线性代数方程组 (因系数矩阵严格对角占优, 方程组可唯一求解), 它所联系的网点分布如图 1.2 (c) 所示.

令

$$L_h^{(3)}u_j^n = \frac{u_j^{n+1}-u_j^n}{\tau} - \frac{a}{2}\left[\frac{u_{j+1}^{n+1}-2u_j^{n+1}+u_{j-1}^{n+1}}{h^2}+\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2}\right].$$

将截断误差

$$R_j^n(u) = L_h^{(3)}u(x_j,t_n)-[Lu]_j^n$$

于 $(x_j,t_{n+\frac{1}{2}})\left(t_{n+\frac{1}{2}}=\left(n+\frac{1}{2}\tau\right)\right)$ 展开, 则得

$$R_j^n(u)=O(\tau^2+h^2). \tag{3.1.9}$$

(四) **Richardson 格式**, 即

$$\frac{u_j^{n+1}-u_j^{n-1}}{2\tau}=a\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2}+f_j, \tag{3.1.10}$$

或

$$u_j^{n+1}=2r(u_{j+1}^n+2u_j^n+u_{j-1}^n)+u_j^{n-1}+2\tau f_j. \tag{3.1.10$'$}$$

这是三层显式差分格式, 它联系的网点分布如图 1.2 (d) 所示. 显然截断误差的阶为 $O(\tau^2+h^2)$. 为了计算能够逐层进行, 除初值 u_j^0 外, 还要用到 u_j^1, 这可以用前述二层差分格式计算 (为保证精度, 可将 $[0,\tau]$ 分成若干等份).

除以上四种差分格式外, 还可做出 (例如用待定系数法, 参看第二章节 2.3) 许多逼近 (3.1.1), (3.1.3) 的差分格式, 但并不是每一差分格式都是可用的. 衡量一个差分格式是否经济实用, 由多方面的因素决定, 主要有:

(1) **计算简单**. 显格式无须解方程组, 计算较隐格式简单. 但隐格式 $(3.1.6)_1'$ 和 $(3.1.8)_1'$ 左端系数是与 n 无关的三对角矩阵, 这相当于求解许多具不同右端但有同一三对角系数矩阵的方程组, 用消元法有许多方便. 再考虑到其他因素 (如稳定性), 隐格式也是可用的.

(2) **收敛性和收敛速度**. 当网比 r 固定, 步长 $h\to 0$ 时, 差分解 u_j^n 应收敛到精确解 $u(x_j,t_n)$, 并希望有尽可能快的收敛速度. 差分算子 L_h 的截断误差的无穷小阶反映了 L_h 对微分算子 L 的逼近程度, 因此可以期望, 截断误差的阶越高, 差分解的精度也越高. 六点对称格式和 Richardson 格式截断误差的阶是 $O(\tau^2+h^2)$, 而显格式 $(3.1.4)_1$ 与隐格式 $(3.1.6)_1$ 是 $O(\tau+h^2)$, 故从截断误差方面看, 六点对称格式和 Richardson 格式有更大优越性.

(3) **稳定性**. 在计算过程中, 由于初始数据有误差, 并且不可避免地有舍入误差, 因此人们自然关心这些误差传递下去, 是无限增长还是可以被控制?这便是稳定性问题. 显然, 只有稳定的差分格式才是可用的.

作为例子, 我们考察 Richardson 格式的稳定性. 由前面知道, Richardson 格式是显格式, 截断误差的阶是 $O(\tau^2+h^2)$. 但从稳定性方面来看, 它是不可用的. 用 e_j^n 表示 u_j^n 的误差, 假定右端 f_j^n 的计算是精确的, 则 e_j^n 满足与 (3.1.10)$'$ 相应的齐方程:

$$e_j^{n+1}=2r(e_{j+1}^n-2e_j^n+e_{j-1}^n)+e_j^{n-1}. \tag{3.1.11}$$

设误差只在初始层的原点 ($j=0$) 发生, 即 $e_j^0=\delta_{j0}\varepsilon(\varepsilon>0;\delta_{00}=1,\delta_{j0}=0$ 当 $j\neq 0), e_j^{-1}=0$, 而在以后计算中都是精确的, 则初始误差的传播如表 1.1.

表 1.1 $r=1/2$ 时 Richardson 格式的误差传播

$k \backslash j$	4	3	–2	–1	0	1	2	3	4
0	0	0	0	0	ε	0	0	0	0
1	0	0	0	ε	-2ε	ε	0	0	0
2	0	0	ε	4ε	7ε	-4ε	ε	0	0
3	0	ε	-6ε	17ε	-24ε	17ε	-6ε	ε	0
4	ε	-8ε	31ε	-68ε	89ε	-68ε	31ε	-8ε	ε
5	-10ε	49ε	-144ε	273ε	-388ε	273ε	-144ε	49ε	-10ε
6	71ε	-260ε	641ε	$-1\,096\varepsilon$	$1\,311\varepsilon$	$-1\,096\varepsilon$	641ε	-260ε	71ε

从表中看出, 误差随 $n\to\infty(h\to 0)$ 无限增长, 所以差分格式不稳定. 表中的计算虽然是就 $r=\dfrac{1}{2}$ 进行的, 实际上对任何 $r>0$ 都有类似现象, 所以 Richardson 格式恒不稳定.

如果采用向前差分格式, 并取 $r=\dfrac{1}{2}$, 则误差方程为

$$e_j^{n+1}=\frac{1}{2}(e_{j+1}^n+e_{j-1}^n). \tag{3.1.12}$$

此时误差逐渐衰减, 如表 1.2 所示. 显然如此的误差传递是允许的. 若限制 $0<r\leqslant\dfrac{1}{2}$, 则误差仍然衰减; 但当 $r>\dfrac{1}{2}$ 时, 误差也无限增长, 所以向前差分格式是条件稳定.

表 1.2 $r=1/2$ 时向前差分格式的误差传播

$k \backslash j$	–4	–3	–2	–1	0	1	2	3	4
0	0	0	0	0	ε	0	0	0	0
1	0	0	0	0.5ε	0	0.5ε	0	0	0
2	0	0	0.25ε	0	0.5ε	0	0.25ε	0	0
3	0	0.125ε	0	0.375ε	0	0.375ε	0	0.125ε	0
4	$0.062\,5\varepsilon$	0	0.25ε	0	0.375ε	0	0.25ε	0	$0.062\,5\varepsilon$

从第一、第二章中还知道, 稳定性不仅对控制误差增长是重要的, 而且也和收敛性有关, 因此稳定性理论在数值求解非驻定问题中占有中心地位.

习 题

1. (实习题) 就 $r=\dfrac{5}{11},\dfrac{5}{9}$ 用显格式 $(3.1.4)'_1$ 作数值实验, 观察误差的增长规律, 并说明 $r=\dfrac{5}{11}$ 时稳定, $r=\dfrac{5}{9}$ 时不稳定.

2. 将向前差分格式和向后差分格式作加权平均, 得到下列格式:

$$\frac{u_j^{n+1}-u_j^n}{\tau}=\frac{a}{h^2}[\theta(u_{j+1}^{n+1}-2u_j^{n+1}+u_{j-1}^{n+1})+(1-\theta)(u_{j+1}^n-2u_j^n+u_{j-1}^n)], \tag{3.1.13}$$

其中 $0 \leqslant \theta \leqslant 1$. 试计算其截断误差, 并证明当 $\theta = \frac{1}{2} - \frac{1}{12r}$ 时, 截断误差的阶最高 $(O(\tau^2 + h^4))$.

3. 在 Richardson 格式 (3.1.10) 中以 $u_j^n = \frac{1}{2}(u_j^{n+1} + u_j^{n-1})$ 代入左端, 便得 Du Fort-Frankel 格式:

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\tau} = a\frac{u_{j+1}^n - u_j^{n+1} - u_j^{n-1} + u_{j-1}^n}{h^2}. \tag{3.1.14}$$

试求其截断误差.

4. 设有逼近热传导方程的带权三层差分格式:

$$(1+\theta)\frac{u_j^{n+1} - u_j^n}{\tau} - \theta\frac{u_j^n - u_j^{n-1}}{\tau} = a\frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{h^2}, \tag{3.1.15}$$

其中 $\theta \geqslant 0$. 试计算其截断误差, 并证明当 $\theta = \frac{1}{2} + \frac{1}{12r}$ 时, 截断误差的阶最高 $(O(\tau^2 + h^4))$.

§2 稳定性与收敛性

2.1 稳定性概念

前节引进的二层差分格式, 均可用矩阵和向量的记号表成

$$\boldsymbol{A}\boldsymbol{U}^{n+1} = \boldsymbol{B}\boldsymbol{U}^n + \tau\boldsymbol{F}, \tag{3.2.1}$$

其中 $\boldsymbol{U}^n = (u_1^n, \cdots, u_{J-1}^n)^{\mathrm{T}}, \boldsymbol{F} = (f_1, \cdots, f_{J-1})^{\mathrm{T}}, \boldsymbol{A}$ 和 $\boldsymbol{B}$ 是 $(J-1)\times(J-1)$ 矩阵. 假定 $\boldsymbol{A}$ 有逆, 并令

$$\boldsymbol{C} = \boldsymbol{A}^{-1}\boldsymbol{B}, \tag{3.2.2}$$

则可将 (3.2.1) 化为

$$\boldsymbol{U}^{n+1} = \boldsymbol{C}\boldsymbol{U}^n + \tau\boldsymbol{A}^{-1}\boldsymbol{F}, \tag{3.2.3}$$

其中 $\boldsymbol{C}$ 称为**增长矩阵**.

例如对于向前差分格式, $\boldsymbol{A} = \boldsymbol{I}$((J-1) 阶单位矩阵), $\boldsymbol{B} = (1-2r)\boldsymbol{I} + r\boldsymbol{S}$, 其中

$$\boldsymbol{S} = \begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & 1 & 0 \end{bmatrix}_{(J-1)\times(J-1)} \tag{3.2.4}$$

故 $\boldsymbol{C}=(1-2r)\boldsymbol{I}+r\boldsymbol{S}$. 对于向后差分格式, $\boldsymbol{A}=(1+2r)\boldsymbol{I}-r\boldsymbol{S}, \boldsymbol{B}=\boldsymbol{I}$ 故 $\boldsymbol{C}=[(1+2r)\boldsymbol{I}-r\boldsymbol{S}]^{-1}$. 对于六点对称格式, $\boldsymbol{A}=(1+r)\boldsymbol{I}-\frac{r}{2}\boldsymbol{S}, \boldsymbol{B}=(1-r)\boldsymbol{I}+\frac{r}{2}\boldsymbol{S}$, 故 $\boldsymbol{C}=\left[(1+r)\boldsymbol{I}-\frac{r}{2}\boldsymbol{S}\right]^{-1}\left[(1-r)+\frac{r}{2}\boldsymbol{S}\right]$.

至于一般的三层或多层格式, 总可适当引进新变量化成二层格式. 例如 Richardson 格式, 其矩阵形式为

$$\boldsymbol{U}^{n+1}=2r(\boldsymbol{S}-2\boldsymbol{I})\boldsymbol{U}^{n}+\boldsymbol{U}^{n-1}. \tag{3.2.5}$$

令 $\boldsymbol{W}^{n}=(\boldsymbol{U}^{n},\boldsymbol{U}^{n-1})^{\mathrm{T}}$, 则化为

$$\boldsymbol{W}^{n+1}=\boldsymbol{C}\boldsymbol{W}^{n}, \tag{3.2.6}$$

其中

$$\boldsymbol{C}=\begin{pmatrix} 2r(\boldsymbol{S}-2\boldsymbol{I}) & \boldsymbol{I} \\ \boldsymbol{I} & \boldsymbol{O} \end{pmatrix}. \tag{3.2.7}$$

我们仅讨论系数及右端与时间 t 无关的方程, 所以 (3.2.1) 的 $\boldsymbol{A},\boldsymbol{B}$ 和 $\boldsymbol{F}$ 均不依赖 n, 但 $\boldsymbol{A},\boldsymbol{B}$ 可依赖步长 h,τ. 我们要求 h,τ 之间满足一定关系, 设为 $h=g(\tau)$, 其中 $g(\tau)$ 连续, $g(0)=0$. 于是 $\boldsymbol{A}=\boldsymbol{A}(\tau),\boldsymbol{B}=\boldsymbol{B}(\tau),\boldsymbol{C}=\boldsymbol{C}(\tau)$.

先讨论按初值稳定. 此时 $\boldsymbol{F}=\boldsymbol{0}$,

$$\boldsymbol{U}^{n+1}=\boldsymbol{C}(\tau)\boldsymbol{U}^{n}=\cdots=[\boldsymbol{C}(\tau)]^{n+1}\boldsymbol{U}^{0}. \tag{3.2.8}$$

我们说差分格式 (3.2.1) **按初值稳定**, 如果存在 $\tau_0>0$ 和常数 $K>0$, 使不等式

$$||\boldsymbol{U}^{n+1}||=||[\boldsymbol{C}(\tau)]^{n+1}\boldsymbol{U}^{0}||\leqslant K||\boldsymbol{U}^{0}|| \tag{3.2.9}$$

对一切 $\boldsymbol{U}^{0}\in\mathbf{R}^{J-1}, 0<\tau\leqslant\tau_0$ 和 $0<n\tau\leqslant T$ 成立. 这里 $||\cdot||$ 是 $\mathbf{R}^{J-1}$ 中的某一种范数, 一般取

$$||\boldsymbol{U}||^2=\|\boldsymbol{u}\|_0^2=\sum_{j=1}^{J-1}hu_j^2.$$

显然差分格式 (3.2.1) 按初值稳定, 当且只当

$$||\boldsymbol{C}^{n}(\tau)||\leqslant K, 0<\tau\leqslant\tau_0, 0<n\tau\leqslant T. \tag{3.2.10}$$

其次讨论按右端稳定. 此时认为初值没有误差, 即 $\boldsymbol{U}^{0}=\boldsymbol{0}$. 我们说差分格式 (3.2.1) **按右端稳定**, 如果存在 $\tau_0>0$ 和常数 $K>0$, 使不等式

$$||\boldsymbol{U}^{n+1}||\leqslant K||\boldsymbol{F}||$$

对一切 $0<\tau\leqslant\tau_0$ 和 $0<n\tau\leqslant T$ 成立, 其中 $\boldsymbol{U}^n$ 是下列方程的解:

(3.2.11) $$\boldsymbol{U}^{n+1}=\boldsymbol{C}(\tau)\boldsymbol{U}^n+\tau\boldsymbol{A}^{-1}\boldsymbol{F},\boldsymbol{U}^0=\boldsymbol{0}.$$

反复利用递推式 (3.2.11), 得

$$\begin{aligned}\boldsymbol{U}^{n+1}&=\boldsymbol{C}(\tau)\boldsymbol{U}^n+\tau\boldsymbol{A}^{-1}\boldsymbol{F}\\&=\boldsymbol{C}(\tau)[\boldsymbol{C}(\tau)\boldsymbol{U}^{n-1}+\tau\boldsymbol{A}^{-1}\boldsymbol{F}]+\tau\boldsymbol{A}^{-1}\boldsymbol{F}\\&=\boldsymbol{C}^2(\tau)\boldsymbol{U}^{n-1}+\tau\boldsymbol{C}(\tau)\boldsymbol{A}^{-1}\boldsymbol{F}+\tau\boldsymbol{A}^{-1}\boldsymbol{F}\\&=\boldsymbol{C}^2(\tau)[\boldsymbol{C}(\tau)\boldsymbol{U}^{n-2}+\tau\boldsymbol{A}^{-1}\boldsymbol{F}]+\tau[\boldsymbol{C}(\tau)+\boldsymbol{I}]\boldsymbol{A}^{-1}\boldsymbol{F}\\&\quad\cdots\\&=\tau\left[\boldsymbol{C}^n(\tau)+\boldsymbol{C}^{n-1}(\tau)+\cdots+\boldsymbol{C}(\tau)+\boldsymbol{I}\right]\boldsymbol{A}^{-1}\boldsymbol{F}.\end{aligned}$$

(注意 $\boldsymbol{U}^0=\boldsymbol{0}$) 设 $\|\boldsymbol{A}^{-1}\|\leqslant K'$, 又差分格式按初值稳定, 即存在常数 K'' 使 $\|\boldsymbol{C}^n(\tau)\|\leqslant K''$, 则

$$\|\boldsymbol{U}^{n+1}\|\leqslant\tau(n+1)K'K''\|\boldsymbol{F}\|\leqslant TK'K''\|\boldsymbol{F}\|.$$

取 $K=TK'K''$, 即知格式按右端稳定.

如果右端与时间有关, 即 $\boldsymbol{F}=\boldsymbol{F}^n$, 以上推导仍成立, 只需用 $\sup\limits_n\|\boldsymbol{F}^n\|$ 代替上述不等式右端的 $\|\boldsymbol{F}\|$. 总之, 若 $\|\boldsymbol{A}^{-1}(\tau)\|\leqslant K'$, 则由格式按初值稳定可推出它按右端稳定. 为检验格式按初值稳定, 需检验不等式 (3.2.10), 即矩阵族

(3.2.12) $$\{\boldsymbol{C}^n(\tau):0<\tau\leqslant\tau_0,0<n\tau\leqslant T\}$$

一致有界. 往后, 我们所述的稳定均指按初值稳定.

2.2 判别稳定性的直接估计法 (矩阵法)

判别矩阵族 (3.2.12) 的一致有界性是一个困难问题, 只在某些特殊情形才能给出解答.

命题 2.1 (必要条件) *以 $\rho(\boldsymbol{C})$ 表示矩阵 $\boldsymbol{C}(\tau)$ 的谱半径, 则差分格式稳定的必要条件是存在与 τ 无关的常数 M 使*

(3.2.13) $$\rho(\boldsymbol{C})\leqslant 1+M\tau\quad(\rho(\boldsymbol{C})\leqslant 1+O(\tau)).$$

证明 由 (3.2.10),

$$\rho^n(\boldsymbol{C})\leqslant\|\boldsymbol{C}^n\|\leqslant K,\quad 0<n\leqslant\frac{T}{\tau},\quad 0<\tau\leqslant\tau_0.$$

不妨设 $K>1$, 并取 $n=\left[\frac{T}{\tau}\right]$ ($\left[\frac{T}{\tau}\right]$ 表示 T/τ 的整数部分), 则

$$\rho(\boldsymbol{C})\leqslant K^{\frac{1}{n}}\leqslant K^{\tau/(T-\tau)}=\mathrm{e}^{\frac{\tau}{T-\tau}\ln K}\leqslant \mathrm{e}^{\frac{\ln K}{T-\tau_0}\tau}=1+O(\tau).$$

命题 2.2 (充分条件) 若 $\boldsymbol{C}(\tau)$ 是正规矩阵, 即 $\boldsymbol{C}$ 和它的共轭转置 $\boldsymbol{C}^*$ 乘积可交换: $\boldsymbol{C}\boldsymbol{C}^*=\boldsymbol{C}^*\boldsymbol{C}$, 则 (3.2.13) 也是差分格式稳定的充分条件.

证明 因为此时 $||\boldsymbol{C}(\tau)||=\rho(\boldsymbol{C})$, 由 (3.2.13),

$$\begin{aligned}||\boldsymbol{C}^n(\tau)||\leqslant||\boldsymbol{C}(\tau)||^n&=\rho^n(\boldsymbol{C})\leqslant(1+M\tau)^n\\&\leqslant(1+M\tau)^{\frac{T}{\tau}}\leqslant K<\infty.\end{aligned}$$

推论 2.1 若 $\boldsymbol{S}$ 是对称矩阵, $\boldsymbol{C}(\tau)$ 是矩阵 $\boldsymbol{S}$ 的实系数有理函数: $\boldsymbol{C}(\tau)=R(\boldsymbol{S})$, 则差分格式稳定的充要条件是

$$\max_j|R(\lambda_j^s)|\leqslant 1+M\tau,$$

其中 λ_j^s 是 $\boldsymbol{S}$ 的特征值. (只需注意 $R(\boldsymbol{S})$ 是实数和矩阵 $\boldsymbol{S}$ 的四则运算)

特别当 $\boldsymbol{S}$ 是形如 (3.2.4) 的矩阵时, 其特征值

$$\lambda_j^S=2\cos j\pi h,\quad j=1,2,\cdots,J-1,\quad h=\frac{l}{J},$$

特征向量 $\boldsymbol{U}^j$ 的分量 $u_k^j=\sin jk\pi h, k=1,2,\cdots,J-1$ (参看第二章 §6 例 6.1 及其习题 1).

例 2.1 对向前差分格式 (以下设 $(3.1.3)_2$ 中的 l=1), $\boldsymbol{C}=(1-2r)\boldsymbol{I}+r\boldsymbol{S}$,

$$\lambda_j^{\boldsymbol{C}}=1-2r+2r\cos j\pi h=1-4r\sin^2\frac{j\pi h}{2}.$$

为使 $|\lambda_j^{\boldsymbol{C}}|\leqslant 1+M\tau$ 或

$$-1-M\tau\leqslant\lambda_j^{\boldsymbol{C}}=1-4r\sin^2\frac{j\pi h}{2}\leqslant 1+M\tau,$$

必须且只需

$$4r\sin^2\frac{j\pi h}{2}\leqslant 2+M\tau,\quad j=1,2,\cdots,J-1,$$

从而 $4r\leqslant 2, r\leqslant\frac{1}{2}$. 所以向前差分格式当 $r\leqslant\frac{1}{2}$ 时稳定, 当 $r>\frac{1}{2}$ 时不稳定.

例 2.2 对向后差分格式, $\boldsymbol{C}=[(1+2r)\boldsymbol{I}-r\boldsymbol{S}]^{-1}$,

$$\lambda_j^{\boldsymbol{C}}=[(1+2r)-2r\cos j\pi h]^{-1}=[1+2r(1-\cos j\pi h)]^{-1}\leqslant 1,$$

故对任何 $r>0$ 稳定, 即**恒稳定**或**绝对稳定**.

例 2.3 对六点对称格式,

$$\boldsymbol{C}=\left[(1+r)\boldsymbol{I}-\frac{r}{2}\boldsymbol{S}\right]^{-1}\left[(1-r)\boldsymbol{I}+\frac{r}{2}\boldsymbol{S}\right],$$

$$\lambda_j^{\boldsymbol{C}}=\frac{1-2r\sin^2\dfrac{j\pi h}{2}}{1+2r\sin^2\dfrac{j\pi h}{2}},\quad j=1,2,\cdots,J-1,$$

故对任何 $r>0$ 有 $|\lambda_j^{\boldsymbol{C}}|\leqslant 1$, 因此六点对称格式恒稳定.

例 2.4 对 Richardson 格式,

$$\boldsymbol{C}=\begin{pmatrix}2r(\boldsymbol{S}-2\boldsymbol{I}) & \boldsymbol{I}\\ \boldsymbol{I} & \boldsymbol{O}\end{pmatrix}$$

是对称矩阵. 设 λ 为 $\boldsymbol{C}$ 的特征值, $\boldsymbol{W}=(\omega_1,\omega_2)^{\mathrm{T}}$ 为相应的特征向量, 即 $\boldsymbol{CW}=\lambda\boldsymbol{W}$, 或

$$2r(\boldsymbol{S}-2\boldsymbol{I})\omega_1+\omega_2=\lambda\omega_1,\quad \omega_1=\lambda\omega_2.$$

显然 $\omega_2\neq 0$. 利用第二方程消去 ω_1, 得

$$2\lambda r(\boldsymbol{S}-2\boldsymbol{I})\omega_2+\omega_2=\lambda^2\omega_2,$$

从而

$$\boldsymbol{S}\omega_2=\left(2+\frac{\lambda}{2r}-\frac{1}{2\lambda r}\right)\omega_2.$$

可见 $\mu=2+\dfrac{\lambda}{2r}-\dfrac{1}{2\lambda r}$ 是 $\boldsymbol{S}$ 的特征值. 于是 λ 满足方程

$$\lambda^2+2r(2-\mu)\lambda-1=0\quad(\mu=2\cos j\pi h),$$

或

$$\lambda^2+\left(8r\sin^2\frac{j\pi h}{2}\right)\lambda-1=0.$$

其根的按模最大值

$$\begin{aligned}\max_j(|\lambda_1^j|,|\lambda_2^j|)&=\max_j\left|4r\sin^2\frac{j\pi h}{2}+\sqrt{16r^2\sin^4\frac{j\pi h}{2}+1}\right|\\&>r+\sqrt{1+r^2}>1+r,\quad 对任意\ r>0,\end{aligned}$$

所以 Richardson 格式恒不稳定.

例 2.5 带第二边值条件的向前差分格式:

$(3.2.14)_1 \qquad u_j^{n+1} = ru_{j+1}^n + (1-2r)u_j^n + u_{j-1}^n, \quad j = 1, 2, \cdots, J-1,$

$(3.2.14)_2 \qquad u_0^n = 0, \quad u_J^n = u_{J-1}^n,$

其中 $u_J^n = u_{J-1}^n$ 是第二边值条件 $u_x(1,t) = 0$ 的差分近似. 利用 $(3.2.14)_2$ 消去 $(3.2.14)_1$ 中的 u_0^n, u_J^n, 则知增长矩阵

$$
\boldsymbol{C} = \begin{bmatrix} 1-2r & r & & & \\ r & 1-2r & r & & \\ & \ddots & \ddots & \ddots & \\ & & r & 1-2r & r \\ & & & r & 1-r \end{bmatrix} = \boldsymbol{I} - r\boldsymbol{D},
$$

其中

$$
\boldsymbol{D} = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix}
$$

是对称矩阵, 其特征值 λ 是实数. 求 $\boldsymbol{D}$ 的特征值等同于解下列差分算子的特征问题:

$$
l_h u_j = -(u_{j+1} + u_{j-1}) = (\lambda - 2)u_j, \quad j = 1, 2, \cdots, J-1, \quad u_0 = 0, \quad u_J = u_{J-1}.
$$

仿照第二章 §6, 以 $u_j = z^j$ 代入上式, 得 $z^2 + (\lambda - 2)z + 1 = 0$, 其解

$$
z_1, z_2 = \frac{1}{2}\left(2 - \lambda \pm \sqrt{\lambda^2 - 4\lambda}\right).
$$

为使解满足左边值条件, 它的二根 z_1, z_2 不能是实的, 而是一对共轭复根 $z_1 = \overline{z}_2 = \mathrm{e}^{\mathrm{i}\theta}$, 且 $\cos\theta = 1 - \dfrac{\lambda}{2}, \sin\theta = \dfrac{\sqrt{4\lambda - \lambda^2}}{2}$. 一般解为 $u_j = c_1 \cos j\theta + c_2 \sin j\theta$. 由边值条件 $u_0 = 0, u_J = u_{J-1}$ 知 $c_1 = 0, \sin J\theta = \sin(J-1)\theta$, 从而 $\theta = (2j-1)\pi/(2J-1)$, $\boldsymbol{D}$ 的特征值为 $\lambda = 2 - 2\cos\theta$, 即

$$
\begin{aligned}
\lambda_j &= 2 - 2\cos\frac{(2j-1)\pi}{2J-1}, \\
&= 4\sin^2\frac{(2j-1)\pi}{2(2J-1)\pi}, \quad j = 1, 2, \cdots, J-1.
\end{aligned}
$$

所以 $\boldsymbol{C}$ 的特征值

$$\lambda_j^{\boldsymbol{C}} = 1 - 4r\sin^2\frac{2j-1}{2(2J-1)}\pi, \quad j = 1, 2, \cdots, J-1$$

为使 $|\lambda_j^{\boldsymbol{C}}| \leqslant 1$ 必须且只需 $r \leqslant \frac{1}{2}$. 又因 $\boldsymbol{C}$ 对称, 故格式 (3.2.14) 的稳定条件是 $r \leqslant \frac{1}{2}$.

2.3 收敛性与敛速估计

如 §1, 考虑热传导方程的初边值问题:

$$\begin{cases} Lu = \dfrac{\partial u}{\partial t} - a\dfrac{\partial^2 u}{\partial x^2} = f(x), \quad 0 < x < l, \quad 0 < t \leqslant T, \\ u(x,0) = \varphi(x), \quad u(0,t) = u(l,t) = 0. \end{cases} \tag{3.2.15}$$

相应的差分格式为

$$\begin{cases} L_h u_j^n = f_j, \quad j = 1, 2, \cdots, J-1, \quad n = 0, 1, \cdots, N-1, \\ u_j^0 = \varphi_j, \quad u_0^n = u_J^n = 0. \end{cases} \tag{3.2.16}$$

其向量形式如 (3.2.3).

差分逼近的截断误差

$$R_j^n(u) = L_h u(x_j, t_n) - [Lu]_j^n, \tag{3.2.17}$$

$u(x,t)$ 是 $0 \leqslant x \leqslant l, 0 \leqslant t \leqslant T$ 上的任一充分光滑函数. 称差分算子 L_h 是边值问题 (3.2.15) 的**相容逼近**, 如果相容条件

$$\lim_{\tau \to 0} ||\mathbf{R}^n|| = 0 \quad (||\mathbf{R}^n|| = o(1)) \tag{3.2.18}$$

成立, 其中 $\mathbf{R}^n$ 是分量为 $R_j^n(u)$ 的向量,$||\cdot||$ 是 $\mathbf{R}^{J-1}$ 中的范数 (参看 (2.1.10)—(2.1.13)).

先对差分解作出某种估计. 将 (3.2.16) 的解分解为 $u_j^n = v_j^n + w_j^n$, 其中 v_j^n 满足零初值和非齐右端方程:

$$\boldsymbol{V}^{n+1} = \boldsymbol{C}(\tau)\boldsymbol{V}^n + \tau\boldsymbol{A}^{-1}\boldsymbol{F} \quad (\boldsymbol{V}^0 = \boldsymbol{0}),$$

而 w_j^n 满足非零初值和齐右端方程:

$$\boldsymbol{W}^{n+1} = \boldsymbol{C}(\tau)\boldsymbol{W}^n \quad (\boldsymbol{W}^0 = \boldsymbol{U}^0),$$

其中 $\boldsymbol{V}^n, \boldsymbol{W}^n$ 依次为以 v_j^n, w_j^n 为分量的向量. 若差分格式按初值稳定, 则亦按右端稳定, 于是有常数 K_1 和 K_2, 使

$$||\boldsymbol{V}^n|| \leqslant K_1||\boldsymbol{F}||, ||\boldsymbol{W}^n|| \leqslant K_2||\boldsymbol{W}^0|| = K_2||\boldsymbol{U}^0||.$$

这样

$$||\boldsymbol{U}^n|| \leqslant K(||\boldsymbol{U}^0|| + ||\boldsymbol{F}||), \quad K = \max(K_1, K_2). \tag{3.2.19}$$

现在估计差分解的误差. 设 $u(x,t)$ 是热传导方程 (3.2.15) 的解, u_j^n 是差分方程 (3.2.16) 的解. 误差 $e_j^n = u(x_j, t_n) - u_j^n = [u]_j^n - u_j^n$. 我们有

$$\begin{aligned} R_j^n(u) &= L_h[u]_j^n - [Lu]_j^n = L_h[u]_j^n - f_j \\ &= L_h[u]_j^n - L_h u_j^n = L_h e_j^n, \end{aligned}$$

即误差 e_j^n 满足差分方程:

$$L_h e_j^n = R_j^n(u), \quad e_j^0 = 0, \quad j = 1, 2, \cdots, J-1.$$

由 (3.2.3) 知其向量形式为

$$\boldsymbol{E}^{n+1} = \boldsymbol{C}(\tau)\boldsymbol{E}^n + \tau \boldsymbol{A}^{-1}\boldsymbol{R}^n,$$

这里 $\boldsymbol{E}^n, \boldsymbol{R}^n$ 依次为以 e_j^n, R_j^n 为分量的向量. 由估计式 (3.2.19) 得

$$||\boldsymbol{E}^n|| \leqslant K \sup_n ||\boldsymbol{R}^n||. \tag{3.2.20}$$

若相容条件 (3.2.18) 成立, 则

$$\lim_{\tau \to 0} ||\boldsymbol{E}^n|| = \lim_{\tau \to 0} ||\boldsymbol{u}^n - \boldsymbol{U}^n|| = 0,$$

其中 $\boldsymbol{u}^n$ 表示以 $u(x_j, t_n)$ 为分量的向量. 这证明了如下

定理 2.1 *若差分方程满足相容条件, 且按初值稳定, 则差分解收敛到热传导方程的解, 且有误差估计式 (3.2.20).*

推论 2.2 *当网比 $r \leqslant \frac{1}{2}$ 时, 向前差分格式的解有收敛阶 $O(\tau + h^2)$. 对任何网比 $r > 0$, 向后差分格式的解有收敛阶 $O(\tau + h^2)$, 六点对称格式的解有收敛阶 $O(\tau^2 + h^2)$.*

[注] 实际上, 定理 2.1 及其证明对更一般的非驻定偏微分方程和差分格式也成立 (参看 [28] 的第 3 章).

习 题

1. 求证差分格式 (3.1.13) 当 $\frac{1}{2} \leqslant \theta \leqslant 1$ 时恒稳定, 当 $0 \leqslant \theta < \frac{1}{2}$ 时稳定的充要条件是

$$r \leqslant 1/2(1-2\theta).$$

2. 证明如下格式恒稳定:

$$\begin{aligned}&\frac{1}{12}\frac{u_{j+1}^{n+1}-u_{j+1}^{n}}{\tau}+\frac{5}{6}\frac{u_{j}^{n+1}-u_{j}^{n}}{\tau}+\frac{1}{12}\frac{u_{j-1}^{n+1}-u_{j-1}^{n}}{\tau}\\&=a\frac{u_{j+1}^{n+1}-2u_{j}^{n+1}+u_{j-1}^{n+1}+u_{j+1}^{n}-2u_{j}^{n}+u_{j-1}^{n}}{2h^2}\quad (a>0).\end{aligned} \tag{3.2.21}$$

§3 Fourier 方 法

前节介绍的判别稳定性的直接估计法, 原则上可用于一般非驻定问题, 但只在某些简单情形才能估计矩阵族 $\boldsymbol{C}^n(\tau)$ 的范数, 这里遇到的主要困难之一是矩阵 $\boldsymbol{C}(\tau)$ 的阶 $(N-1)$ 随 $\tau\to 0$ 而无限增大. 本节仅限于讨论常系数线性非驻定方程的纯初值问题和带周期边值条件的混合问题, 此时可用 Fourier 方法 (Fourier 积分和 Fourier 级数) 将空间变量和时间变量分离, 从而将偏微分方程化为常微分方程, 然后再讨论解的适定性. 本节将 Fourier 方法用于相应的差分方程, 得到若干便于应用的判别稳定性的代数准则.

先考虑线性常系数一维抛物型方程, 具初值和周期 (设周期为 l) 边值条件. 逼近它的二层差分方程的一般形式为

$$\sum_{m\in\mathcal{N}_1} a_m u_{j+m}^{n+1} = \sum_{m\in\mathcal{N}_0} b_m u_{j+m}^{n},\quad j=0,1,\cdots,J-1 \tag{3.3.1}$$

(只考虑按初值稳定, 故可设非齐项等于零). 这是在空间网点 x_j 处的差分方程, $\mathcal{N}_0$ 和 $\mathcal{N}_1$ 是包含 0 及其附近的正负整数的有限集合, a_m 和 b_m 不依赖 j 但可能和 τ 有关. 例如对向前差分格式

$$u_j^{n+1} = ru_{j+1}^n + (1-2r)u_j^n + ru_{j-1}^n,$$

$$\mathcal{N}_0=\{-1,0,1\},\quad \mathcal{N}_1=\{0\},\quad b_{-1}=b_1=r,\quad b_0=1-2r,\quad a_0=1.$$

对向后差分格式

$$-ru_{j+1}^{n+1}+(1+2r)u_j^{n+1}-ru_{j-1}^{n+1}=u_j^n,$$

$$\mathcal{N}_0=\{0\},\quad \mathcal{N}_1=\{-1,0,1\},\quad b_0=1,\quad a_{-1}=a_1=-r,\quad a_0=1+2r.$$

对六点对称格式

$$-\frac{r}{2}u_{j+1}^{n+1}+(1+r)u_j^{n+1}-\frac{r}{2}u_{j-1}^{n+1}=\frac{r}{2}u_{j+1}^{n}+(1-r)u_j^{n}+\frac{r}{2}u_{j-1}^{n},$$

$$\mathscr{N}_0=\mathscr{N}_1=\{-1,0,1\},\quad b_{-1}=b_1=\frac{r}{2},\quad b_0=1-r,\quad a_{-1}=a_1=-\frac{r}{2},\quad a_0=1+r.$$

由于是周期边值条件 ($u_0^n = u_J^n$), 故可将 u_j^n 周期开拓使其对一切 $j = 0, \pm1, \cdots$ 有意义, 且方程 (3.3.1) 对所有整数 j 成立. 为了应用 Fourier 方法, 我们再将 $u_j^n = u^n(x_j)$ 开拓为 $(-\infty, \infty)$ 上的 $u^n(x)$. 为此, 取半整数点 $x_{j+\frac{1}{2}} = x_j + \frac{1}{2}h, j = 0, \pm1, \cdots$, 并用如下阶梯函数逼近初始函数 $\phi(x)$:

$$u^0(x)=\varphi(x_j),\ \text{当}\ x_{j-\frac{1}{2}}<x<x_{j+\frac{1}{2}},$$

其中 $j = 0, \pm1, \cdots$. 再将 (3.3.1) 看成在任一 $x_j = x \in (-\infty, \infty)$ 成立, 则得具连续变量的差分解 $u^n(x)$. 显然 $u^n(x)$ 仍是 x 的周期函数 (周期为 l), 且

$$u^n(x)=u_j^n,\ \text{当} x_{j-\frac{1}{2}}<x<x_{j+\frac{1}{2}},$$

其中 $j = 0, \pm1, \cdots$. 显然 $u^n(x)$ 于 $(0,l)$ 平方可积, 因此属于空间 $L^2(0,l)$, 其范数

$$\|u^n\|_{L^2}^2=\int_0^l |u^n(x)|^2\mathrm{d}x.$$

此外还有

$$\begin{aligned}||U^n||^2&=\sum_{j=0}^{J-1}h(u_j^n)^2=\sum_{j=1}^{J-1}h(u_j^n)^2+\frac{h}{2}\{(u_0^n)^2+(u_N^n)^2\}\\&=\int_0^l|u^n(x)|^2\mathrm{d}x=||u^n(x)||_{L^2}^2.\end{aligned}$$

这样, 我们就可将 Fourier 方法用于具连续空间变量的差分方程

$$\sum_{m\in\mathscr{N}_1}a_m u^{n+1}(x+x_m)=\sum_{m\in\mathscr{N}_0}b_m u^n(x+x_m). \tag{3.3.2}$$

将 $u^n(x)$ 展成 Fourier 级数:

$$u^n(x)=\sum_{p=-\infty}^{\infty}v_p^n\exp\left(\mathrm{i}\frac{2p\pi}{l}x\right), \tag{3.3.3}$$

$$v_p^n=\frac{1}{l}\int_0^l u^n(x)\exp\left(-\mathrm{i}\frac{2p\pi}{l}x\right)\mathrm{d}x,\quad p=0,\pm1,\cdots. \tag{3.3.4}$$

我们有 Parseval 等式:

$$||u^n(x)||_{L^2}^2=l\sum_{p=-\infty}^{\infty}|v_p^n|^2. \tag{3.3.5}$$

把 (3.3.3) 代到 (3.3.2), 得

$$\sum_{p=-\infty}^{\infty} v_p^{n+1}\left[\sum_{m\in\mathcal{N}_1} a_m \exp\left(\mathrm{i}\frac{2p\pi}{l}x_m\right)\right]\exp\left(\mathrm{i}\frac{2p\pi}{l}x\right) = \sum_{p=-\infty}^{\infty} v_p^{n}\left[\sum_{m\in\mathcal{N}_0} b_m \exp\left(\mathrm{i}\frac{2p\pi}{l}x_m\right)\right]\exp\left(\mathrm{i}\frac{2p\pi}{l}x\right). \tag{3.3.6}$$

比较对应项的系数, 得

$$v_p^{n+1} = G(ph,\tau)v_p^n, \tag{3.3.7}$$

其中

$$G(ph,\tau) = \left[\sum_{m\in\mathcal{N}_1} a_m \exp\left(\mathrm{i}\frac{2m\pi}{l}ph\right)\right]^{-1}\left[\sum_{m\in\mathcal{N}_0} b_m \exp\left(\mathrm{i}\frac{2m\pi}{l}ph\right)\right]. \tag{3.3.8}$$

将 (3.3.7) 代到 (3.3.5), 则

$$||u^n(x)||_{L^2}^2 = l\sum_{p=-\infty}^{\infty}|G(ph,\tau)v_p^{n-1}|^2 = l\sum_{p=-\infty}^{\infty}|G^n(ph,\tau)v_p^0|^2. \tag{3.3.9}$$

若差分格式稳定, 则有常数 $K>0$ 使

$$||u^n(x)||_{L^2}^2 = l\sum_{p=-\infty}^{\infty}|G^n(ph,\tau)v_p^0|^2 \leqslant lK^2||u^0(x)||_{L^2}^2.$$

由于阶梯函数类 $\{u^0(x)\}$ 于 $L^2(0,l)$ 稠密, 取 $u^0(x)$ 使 $||u^0(x)||_{L^2}=1$, 其 Fourier 系数 $v_p^0=1, v_q^0=0$ 当 $q\neq p$. 则由上式得

$$|G^n(ph,\tau)|\leqslant K,\quad 0<\tau\leqslant\tau_0,\quad 0<n\tau\leqslant T, \tag{3.3.10}$$

即 $G^n(ph,\tau)$ 一致有界. 反之, 若 $G^n(ph,\tau)$ 一致有界, 则由 (3.3.9) 得

$$||u^n(x)||_{L^2}^2 \leqslant K^2 l\sum_{p=-\infty}^{\infty}|v_p^0|^2 = K^2||u^0(x)||_{L^2}^2, \tag{3.3.11}$$

从而差分格式按初值稳定.

往后我们称 $G(ph,\tau)$ 为**增长因子** (Amplification factor). 不妨设 $K\geqslant 1$. 显然不等式 (3.3.10) 又等价于 $|G(ph,\tau)|\leqslant K^{\frac{1}{n}}$, 取 $n=\left[\frac{T}{\tau}\right]$ (取整数部分), 则 $\frac{1}{n}\leqslant\frac{\tau}{T-\tau}$, 所以 $|G(ph,\tau)|\leqslant K^{\frac{\tau}{T-\tau}}\leqslant\exp\left(\frac{\tau}{T-\tau}\ln K\right)$, 于是

$$|G(ph,\tau)|\leqslant 1+M\tau. \tag{3.3.12}$$

上式也称为 von Neumann 条件. 综合上述, 我们得

命题 3.1　差分格式 (3.3.1) 稳定 $\Longleftrightarrow G^n(ph,\tau)$ 一致有界 $\Longleftrightarrow$ von Neumann 条件 (3.3.12) 成立.

注 3.1　注意 (3.3.8) 中 $ph=x_p$ 是空间网点, $G(x_p,\tau)$ 关于 x_p,τ 连续, 关于 x_p 是以 l 为周期的函数, 所以只需就 $p=0,1,\cdots,J-1$ 研究 $G^n(x_p,\tau)$ 的一致有界性, 此时 $x_0=0<x_1<\cdots<x_{J-1}<l$.

注 3.2　增长因子的计算. 以 Fourier 展式 (3.3.3) 的通项

$$v^n\cdot\exp(\mathrm{i}\alpha x_j)\quad(\alpha=2p\pi/l) \tag{3.3.13}$$

代到 (3.3.1) 两端, 得

$$v^{n+1}\sum_{m\in\mathcal{N}_1}a_m\exp(\mathrm{i}\alpha x_{m+j})=v^n\sum_{m\in\mathcal{N}_0}b_m\exp(\mathrm{i}\alpha x_{m+j}).$$

消去公因子 $\mathrm{e}^{\mathrm{i}\alpha x_j}$, 即得

$$v^{n+1}=\left[\sum_{m\in\mathcal{N}_1}a_m\exp(\mathrm{i}\alpha x_m)\right]^{-1}\left[\sum_{m\in\mathcal{N}_0}b_m\exp(\mathrm{i}\alpha x_m)\right]v^n.$$

v^n 前面的因子就是增长因子 $G(ph,\tau)$ (参看 (3.3.8)).

注 3.3　如果求解的是纯初值问题, 则需用 Fourier 积分

$$u^n(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}v^n(s)\mathrm{e}^{\mathrm{i}xs}\mathrm{d}s \tag{3.3.14}$$

代替 Fourier 级数 (3.3.3). 为使 (3.3.14) 有意义, 应要求初值 $\varphi(x)\in L^2(-\infty,\infty)$. 计算增长因子的方法如注 3.2, 不再详述.

例 3.1　考虑向前差分格式

$$u_j^{n+1}=ru_{j+1}^n+(1-2r)u_j^n+ru_{j-1}^n,$$

以 $u_j^n=v^n\mathrm{e}^{\mathrm{i}\alpha jh}$ 代入, 得

$$v^{n+1}\mathrm{e}^{\mathrm{i}\alpha jh}=(r\mathrm{e}^{\mathrm{i}\alpha(j+1)h}+(1-2r)\mathrm{e}^{\mathrm{i}\alpha jh}+r\mathrm{e}^{\mathrm{i}\alpha(j-1)h})v^n.$$

消去 $e^{\mathrm{i}\alpha jh}$, 则知增长因子

$$\begin{aligned}G(x_p,\tau)&=(1-2r)+r(\mathrm{e}^{\mathrm{i}\alpha h}+\mathrm{e}^{-\mathrm{i}\alpha h})\\&=1-2r(1-\cos\alpha h)\\&=1-4r\sin^2\frac{\alpha h}{2}.\end{aligned}$$

由于 $\alpha h/2\left(=\dfrac{\pi ph}{l}\right)$ 在 $[0,\pi]$ 中分布稠密 (随 $h\to 0$), 为使 $G(x_p,\tau)$ 满足 von Neumann 条件, 必须且只需网比 $r\leqslant\dfrac{1}{2}$ (见 §2 例 2.1), 所以向前差分格式的稳定条件是 $r\leqslant\dfrac{1}{2}$.

注 3.4 Fourier 方法同样可以分析差分方程组的稳定性. 设差分方程组形如

$$\sum_{m\in N_1}\boldsymbol{A}_m\boldsymbol{U}_{j+m}^{n+1}=\sum_{m\in N_0}\boldsymbol{B}_m\boldsymbol{U}_{j+m}^{n},\tag{3.3.15}$$

其中 $\boldsymbol{A}_m,\boldsymbol{B}_m$ 是 $s\times s$ 方阵, 一般依赖步长 τ, 但和 j 无关; $\boldsymbol{U}_j^n$ 是 s 维列向量, 其分量为 $u_{1j}^n,\cdots,u_{sj}^n$. 像方程式的情形一样, 将 $\boldsymbol{U}_j^n$ 开拓为连续变量的周期函数 $\boldsymbol{U}^n(x)=(u_1^n(x),\cdots,u_s^n(x))^{\mathrm{T}}$, 并将它展成 Fourier 级数

$$\boldsymbol{U}^n(x)=\sum_{p=-\infty}^{\infty}\boldsymbol{V}_p^n\exp\left(\mathrm{i}\frac{2p\pi}{l}x\right),$$

将其代入 (3.3.15), 比较相应项的系数, 则得增长因子:

$$\boldsymbol{G}(x_p,\tau)=\left[\sum_{m\in\mathcal{N}_1}\boldsymbol{A}_m\exp\left(\mathrm{i}\frac{2m\pi}{l}x_p\right)\right]^{-1}\left[\sum_{m\in\mathcal{N}_0}\boldsymbol{B}_m\exp\left(\mathrm{i}\frac{2m\pi}{l}x_p\right)\right].\tag{3.3.16}$$

此时 $\boldsymbol{G}(x_p,\tau)$ 是矩阵, 称为增长矩阵. 计算增长矩阵的方法跟以前一样, 以通项

$$\boldsymbol{V}^n\exp(\mathrm{i}\alpha x_j)\quad(\alpha=2p\pi/l)\tag{3.3.17}$$

代到方程 (3.3.15), 消去共同因子 $\exp(\mathrm{i}\alpha x_j)$, 则得

$$\boldsymbol{V}^{n+1}=\boldsymbol{G}(x_p,\tau)\boldsymbol{V}^n,\tag{3.3.18}$$

其中 $\boldsymbol{G}(x_p,\tau)$ 就是形如 (3.3.16) 的增长矩阵.

与前面类似地有

命题 3.2 差分格式 (3.3.15) 稳定的充要条件是: 矩阵族

$$\{\boldsymbol{G}^n(x_p,\tau):0<\tau\leqslant\tau_0,0<n\tau\leqslant T,p=0,1,\cdots,J-1\}\tag{3.3.19}$$

一致有界.

命题 3.3 矩阵族 (3.3.19) 一致有界的必要条件是 $\boldsymbol{G}(x_p,\tau)$ 的谱半径

$$\rho(\boldsymbol{G})\leqslant 1+O(\tau),\tag{3.3.20}$$

即 von Neumann 条件成立.

例 3.2 将 Richardson 格式写成等价的方程组:

$$\begin{cases} u_j^{n+1} = 2r(u_{j+1}^n - 2u_j^n + u_{j-1}^n) + w_j^n, \\ w_j^{n+1} = u_j^n. \end{cases} \tag{3.3.21}$$

以 $u_j^n = v_1^n \mathrm{e}^{\mathrm{i}\alpha x_j}, w_j^n = v_2^n \mathrm{e}^{\mathrm{i}\alpha x_j}$ 代入, 并消去公因子, 得

$$\begin{cases} v_1^{n+1} = 4r(\cos\alpha h - 1)v_1^n + v_2^n, \\ v_2^{n+1} = v_1^n. \end{cases} \tag{3.3.21$'$}$$

显然增长矩阵

$$\boldsymbol{G}(\alpha h) = \begin{pmatrix} -8r\sin^2\dfrac{\alpha h}{2} & 1 \\ 1 & 0 \end{pmatrix}.$$

由 §2 例 2.4 的计算结果, 知 $\boldsymbol{G}(\alpha h)$ 的谱半径对 $\forall r > 0$ 不满足 von Neumann 条件, 故 Richardson 格式恒不稳定.

注 3.5 Fourier 方法也可用于多维差分格式, 求解域或为全空间 (纯初值问题), 或为超长方体 (周期边值条件). 作为例子, 考虑二维热传导方程的初边值问题

$$\begin{cases} \dfrac{\partial u}{\partial t} = a\left(\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2}\right), \quad 0 < x, y < l \ (a > 0), \\ u(x,y,0) = \varphi(x,y), \\ u(0,y,t) = u(l,y,t) = 0, \quad u(x,0,t) = u(x,l,t) = 0. \end{cases} \tag{3.3.22}$$

取步长 $h = l/J, \tau = T/N$, 用两族平行线 $x = x_j = jh, y = y_k = kh$ 将求解域分划成矩形网格, 网点为 $(x_j, y_k, t_n)(t_n = n\tau)$. 引进二阶差分算子

$$\delta_x^2 u_{jk}^n = u_{j+1,k}^n - 2u_{j,k}^n + u_{j-1,k}^n,$$
$$\delta_y^2 u_{jk}^n = u_{j,k+1}^n - 2u_{j,k}^n + u_{j,k-1}^n,$$

作逼近 (3.3.22) 的向前差分格式

$$\frac{u_{j,k}^{n+1} - u_{j,k}^n}{\tau} = \frac{a}{h^2}(\delta_x^2 u_{jk}^n + \delta_y^2 u_{jk}^n), \tag{3.3.23}$$

向后差分格式

$$\frac{u_{j,k}^{n+1} - u_{j,k}^n}{\tau} = \frac{a}{h^2}(\delta_x^2 u_{jk}^{n+1} + \delta_y^2 u_{jk}^{n+1}) \tag{3.3.24}$$

和 Crank–Nicolson 格式

$$\frac{u_{j,k}^{n+1} - u_{j,k}^n}{\tau} = \frac{a}{2h^2}(\delta_x^2 u_{jk}^{n+1} + \delta_x^2 u_{jk}^n + \delta_y^2 u_{jk}^{n+1} + \delta_y^2 u_{jk}^n). \tag{3.3.25}$$

它们的截断误差的阶依次为 $O(\tau+h^2), O(\tau+h^2)$ 和 $O(\tau^2+h^2)$.

现在研究 (3.3.23) 的稳定性. 取通项

$$u_{jk}^n = v^n \exp\left(\mathrm{i}(\alpha x_j + \beta y_k)\right), \quad \alpha = \frac{2\pi p}{l}, \quad \beta = \frac{2\pi q}{l}$$

(参看 (3.3.13)), 代到 (3.3.23) 两端并消去公因子, 得

$$v^{n+1} = \left(1 - 4r\sin^2\frac{\alpha h}{2} - 4r\sin^2\frac{\beta h}{2}\right)v^n \quad (r = a\tau/h^2),$$

从而增长因子

$$G = G(\alpha h, \beta h) = 1 - 4r\left(\sin^2\frac{\alpha h}{2} + \sin^2\frac{\beta h}{2}\right).$$

为使 $|G| = 1 + O(\tau)$, 必须且只需 $r \leqslant \frac{1}{4}$. 由此可见, 随着维数的增加, 对网比的限制更严了.

用同样的方法可证隐格式 (3.3.24) 和 (3.3.25) 恒稳定. 逐层计算需要求解形如

$$u_{jk}^{n+1} - c(\delta_x^2 u_{jk}^{n+1} + \delta_y^2 u_{jk}^{n+1}) = f(u_{jk}^n) \tag{3.3.26}$$

的方程 (对 (3.3.24), $c = r$; 对 (3.3.25), $c = \frac{r}{2}$), 虽然第二章的方法仍可以采用, 但计算量明显增加了.

习　　题

1. 用 Fourier 方法给出差分格式 (3.1.13) 的稳定条件.
2. 证明格式 (3.1.15) 恒稳定.
3. 证明格式 (3.3.24) 和 (3.3.25) 恒稳定.

§4　判别差分格式稳定性的代数准则

由命题 3.2 知道, 差分格式的稳定性归结为矩阵族 (3.3.19) 的一致有界性. 当增长矩阵 $\boldsymbol{G}$ 的阶 $s=1$ 时, $\boldsymbol{G}^n$ 一致有界的充要条件是 von Neumann 条件成立, 即 $|\boldsymbol{G}| = 1 + O(\tau)$. 当 $s \geqslant 2$ 时 von Neumann 条件只是稳定的必要条件, 不是充分条件. 但若 $\boldsymbol{G}(x_p, \tau)$ 是正规矩阵 (特别是对称矩阵), 则 $\boldsymbol{G}(x_p, \tau)$ 的欧氏模等于谱半径, 于是 von Neumann 条件也是稳定的充分条件. 然而增长矩阵一般不是正规矩阵, 因此有必要寻求新的充分条件.

定理 4.1 设 $\boldsymbol{G}(x_p,\tau)$ 关于 τ 于 $\tau=0$ 满足 Lpschitz 条件, 则矩阵族 (3.3.19) 一致有界的充要条件是矩阵族

$$\{\boldsymbol{G}^n(x_p,0): 0<\tau\leqslant\tau_0,\quad 0<n\tau\leqslant T,\quad p=0,1,\cdots,J-1\}$$

一致有界 (见 [14]).

证明 只证充分性, 因为必要性的证明完全类似. 由假设, 有常数 K 使 $\boldsymbol{G}(x_p,\tau)-\boldsymbol{G}(x_p,0)=\tau\boldsymbol{G}_1$, 而 $||\boldsymbol{G}_1||\leqslant K$. 记 $\boldsymbol{G}_0=\boldsymbol{G}(x_p,0)$, 则 $\boldsymbol{G}=\boldsymbol{G}(x_p,\tau)=\boldsymbol{G}_0+\tau\boldsymbol{G}_1$.

注意

$$\begin{aligned}\boldsymbol{G}^n&=(\boldsymbol{G}_0+\tau\boldsymbol{G}_1)\boldsymbol{G}^{n-1}=\boldsymbol{G}_0\boldsymbol{G}^{n-1}+\tau\boldsymbol{G}_1\boldsymbol{G}^{n-1}\\&=\boldsymbol{G}_0(\boldsymbol{G}_0+\tau\boldsymbol{G}_1)\boldsymbol{G}^{n-2}+\tau\boldsymbol{G}_1\boldsymbol{G}^{n-1}\\&=\boldsymbol{G}_0^2\boldsymbol{G}^{n-2}+\tau\boldsymbol{G}_0\boldsymbol{G}_1\boldsymbol{G}^{n-2}+\tau\boldsymbol{G}_1\boldsymbol{G}^{n-1}\\&\quad\cdots\\&=\boldsymbol{G}_0^n+\tau\sum_{i=0}^{n-1}\boldsymbol{G}_0^i\boldsymbol{G}_1\boldsymbol{G}^{n-i-1}.\end{aligned}$$

由于 $\boldsymbol{G}_0^n$ 一致有界, 可设 $||\boldsymbol{G}_0^i||\leqslant M, ||\boldsymbol{G}_0^i\boldsymbol{G}_1||\leqslant M$, 从而得

$$\begin{aligned}||\boldsymbol{G}^n||&\leqslant||\boldsymbol{G}_0^n||+\tau\sum_{i=0}^{n-1}||\boldsymbol{G}_0^i\boldsymbol{G}_1||||\boldsymbol{G}^{n-i-1}||\\&\leqslant M\left[1+\tau\sum_{i=0}^{n-1}||G^{n-i-1}||\right].\end{aligned}$$

由 Gronwall 不等式 (见第一章 §1 引理 1.3), 得

$$||\boldsymbol{G}^n||\leqslant M(1+M\tau)^{n-1}.$$

又 $0<n\leqslant T/\tau$, 故不等式右端一致有界.

例 4.1 考虑逼近带低阶项的抛物方程:

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}+bu.$$

逼近它的的向前差分格式为

$$\frac{u_j^{n+1}-u_j^n}{\tau}=\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2}+bu_j^n.$$

用 Fourier 方法即知增长因子 $G = 1 - 4r\sin^2\dfrac{\alpha h}{2} + b\tau$. 由定理 4.1, G^n 一致有界等价于 $G_0^n \left(G_0 = 1 - 4r\sin^2\dfrac{\alpha h}{2}\right)$ 一致有界. 由 §3 例 3.1, G_0^n 一致有界的充要条件是 $r \leqslant \dfrac{1}{2}$. 这说明低阶项 bu_j^n 不影响差分格式的稳定性.

今设 $\boldsymbol{G}(x_p, \tau) = \boldsymbol{G}(x_p)$ 与 τ 无关. 考虑矩阵族

$$\{\boldsymbol{G}^n(x_p); x_0 = 0 < x_1 < \cdots < x_J = l, \quad 0 < \tau \leqslant \tau_0, \quad 0 < n\tau \leqslant T\} \tag{3.4.1}$$

的一致有界性.

命题 4.1 矩阵族 (3.4.1) 一致有界的充要条件是矩阵族

$$\{\boldsymbol{G}^n(x); 0 \leqslant x \leqslant l, \quad k = 1, 2, \cdots\} \tag{3.4.2}$$

一致有界.

证明 充分性显然, 只证必要性. 假定网格按 2 等分, 4 等分,$\cdots\cdots$, 2^m 等分加密, 则二等分点 x_j 一旦是 $[0, l]$ 的网点便永远是网点. 由假设,

$$\|\boldsymbol{G}^n(x_j)\| \leqslant M, \quad 0 < n\tau \leqslant T,$$

M 是与分划无关的常数. 令 $\tau \to 0$ (从而 $h \to 0$), 则

$$\|\boldsymbol{G}^n(x_j)\| \leqslant M, \quad n = 1, 2, \cdots,$$

而二等分点 $\{x_j\}$ 于 $[0, l]$ 稠密, $G(x)$ 是连续函数, 故

$$\|G^n(x)\| \leqslant M, \quad 0 \leqslant x \leqslant l, \quad n = 1, 2, \cdots.$$

定理 4.2 (一致对角化) 若对任一 $\boldsymbol{G}(x_p, \tau)$, 有矩阵 $\boldsymbol{H}$ 使

$$\boldsymbol{H}^{-1}\boldsymbol{G}H = \boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \ddots \\ & & & \lambda_s \end{bmatrix},$$

且 $\boldsymbol{H}$ 和 $\boldsymbol{H}^{-1}$ 关于 p 和充分小的 $\tau > 0$ 一致有界, 则 von Neumann 条件也是稳定的充分条件.

证明 因 $\boldsymbol{G}^n = \boldsymbol{H}\boldsymbol{\Lambda}^n\boldsymbol{H}^{-1}$, 故结论显然.

若差分方程组 (2.3.15) 是二阶方程组, 则增长矩阵 $\boldsymbol{G}(x, \tau)$ 也是二阶, 此时有一个便于检验的稳定性条件.

定理 4.3　设 $\boldsymbol{G}(x,\tau)$ 是二阶矩阵, g_{ij} 是 $\boldsymbol{G}$ 的第 i 行第 j 列元素, λ_1 和 λ_2 是 $\boldsymbol{G}$ 的特征值, 若下列条件成立:

$$(\alpha)\ \ |\lambda_i(x,\tau)| \leqslant 1 + M\tau,\quad i=1,2,$$

$$(\beta)\ \ \left\|\boldsymbol{G}(x,\tau) - \frac{1}{2}(g_{11}(x,\tau)+g_{22}(x,\tau))\boldsymbol{I}\right\| \leqslant M(\tau + |1-|\lambda_1(x,\tau)|| + |\lambda_1(x,\tau)-\lambda_2(x,\tau)|),$$

则矩阵族 (3.3.19) 一致有界, 其中 $\boldsymbol{I}$ 是二阶单位矩阵.

证明　注意条件 (β) 实际上是

$$\|\boldsymbol{G} - \frac{1}{2}(\lambda_1+\lambda_2)\boldsymbol{I}\| \leqslant M(\tau + |1-|\lambda_1|| + |\lambda_1-\lambda_2|). \tag{3.4.3}$$

利用 $\boldsymbol{G}$ 的 Jordan 标准形, 我们有

$$\boldsymbol{G}^n - \lambda^n\boldsymbol{I} = n\lambda^{n-1}(\boldsymbol{G}-\lambda\boldsymbol{I}),\quad \text{当 } \lambda_1=\lambda_2=\lambda, \tag{3.4.4}$$

$$\boldsymbol{G}^n - \frac{1}{2}(\lambda_1^n+\lambda_2^n)\boldsymbol{I} = \frac{\lambda_1^n-\lambda_2^n}{\lambda_1-\lambda_2}\left(\boldsymbol{G}-\frac{1}{2}(\lambda_1+\lambda_2)\boldsymbol{I}\right),\ \text{当 } \lambda_1\neq\lambda_2. \tag{3.4.5}$$

形式上也可将 (3.4.4) 写成 (3.4.5). 设 $\lambda_1 = r_1\mathrm{e}^{\mathrm{i}\theta_1}, \lambda_2 = r_2\mathrm{e}^{\mathrm{i}\theta_2}$, 不妨设 $r_1 \geqslant r_2$.

由条件 (α), λ_1^n 和 λ_2^n 一致有界. 由 (3.4.5), 可将 $\boldsymbol{G}^n$ 一致有界归结为 (3.4.5) 右端一致有界. 因 $|\lambda_1^n-\lambda_2^n|/|\lambda_1-\lambda_2| \leqslant nr_1^{n-1}, n\tau\leqslant T$, 由 (3.4.3) 可知只需证明 $nr_1^{n-1}|1-r_1|$ 一致有界. 当 $r_1<1$ 时, $nr_1^{n-1}(1-r_1) \leqslant (1+r_1+\cdots+r_1^{n-1})(1-r_1) = 1-r_1^n \leqslant 1$; 当 $r_1\geqslant 1$ 时由 (α) 知 $|1-r_1|\leqslant M\tau$, 从而 $nr_1^{n-1}|1-r_1| \leqslant Mn\tau \leqslant M_T$.

注 4.1　可以证明, $(\alpha),(\beta)$ 也是矩阵族 (3.3.19) 一致有界的必要条件 (参看 [9]).

推论 4.1　特别若 $\boldsymbol{G}(x,\tau)$ 与 τ 无关, 则知二阶矩阵族 (3.4.2) 一致有界的充要条件是

$$(\alpha')\ \ |\lambda_1(x)| \leqslant 1,\quad i=1,2,\quad 0\leqslant x\leqslant l,$$

$$(\beta')\ \ \left\|\boldsymbol{G}(x) - \frac{1}{2}(g_{11}(x)+g_{22}(x))\boldsymbol{I}\right\| \leqslant M(|1-|\lambda_1(x)|| + |\lambda_1(x)-\lambda_2(x)|),\quad 0\leqslant x\leqslant l.$$

注 4.2　定理 4.3 是一个很有用的稳定性判别法, 应用时注意到以下两点是方便的:

1. 实系数二次方程 $\lambda^2 - b\lambda - c = 0$ 的根按模不大于 1 的充要条件是

$$|b| \leqslant 1-c \leqslant 2. \tag{3.4.6}$$

这在检验 von Neumann 条件时有用.

2. 检验条件 (β) 时要计算二阶矩阵的范数, 通常用 Frobinius 范数 $\|\cdot\|_F$ (F−范数). m 阶矩阵 $\boldsymbol{A}=(a_{ij})$ 的 F− 范数是

$$\|\boldsymbol{A}\|_F=\left(\sum_{i,j=1}^{m}|a_{ij}|^2\right)^{\frac{1}{2}}. \tag{3.4.7}$$

例 4.2 考虑逼近热传导方程的 **Dufort–Frankel 格式** (见 §1 习题 3):

$$\frac{u_j^{n+1}-u_j^n}{2\tau}=a\frac{u_{j+1}^n-u_j^{n+1}-u_j^{n-1}+u_{j-1}^n}{h^2}. \tag{3.4.8}$$

引进新变量 $v_j^{n+1}=u_j^n$, 将它化为一阶方程组:

$$\begin{cases} u_j^{n+1}=\dfrac{2r}{1+2r}(u_{j+1}^n+u_{j-1}^n)+\dfrac{1-2r}{1+2r}v_j^n, \\ v_j^{n+1}=u_j^n, \quad r=\tau/h^2 \end{cases}$$

或

$$\begin{pmatrix} u_j^{n+1} \\ v_j^{n+1}\end{pmatrix}=\begin{bmatrix}\dfrac{2r}{1+2r}(T_1+T_{-1}) & \dfrac{1-2r}{1+2r} \\ 1 & 0\end{bmatrix}\begin{pmatrix} u_j^n \\ v_j^n\end{pmatrix},$$

其中 $T_{\pm1}u_j=u_{j\pm1}$ 为移位算子. 用 Fourier 方法可知增长矩阵

$$\boldsymbol{G}(\alpha h)=\begin{bmatrix}\dfrac{2r}{1+2r}(e^{i\alpha h}+e^{-i\alpha h}) & \dfrac{1-2r}{1+2r} \\ 1 & 0\end{bmatrix}=\begin{bmatrix}\dfrac{4r}{1+2r}\cos\alpha h & \dfrac{1-2r}{1+2r} \\ 1 & 0\end{bmatrix}.$$

其特征方程为

$$\lambda^2-\frac{4r\cos\theta}{1+2r}\lambda-\frac{1-2r}{1+2r}=0, \quad \theta=\alpha h=\frac{2\pi ph}{l}\in[0,2\pi], \tag{3.4.9}$$

它的系数显然满足条件 (3.4.6), 故其特征值按模 $\leqslant 1$, 从而条件 $(\alpha)'$ 成立。

其次, (3.4.9) 的二根为

$$\lambda_{1,2}=\left(2r\cos\theta\pm\sqrt{1-4r^2\sin^2\theta}\right)\Big/(1+2r). \tag{3.4.10}$$

令 $\Lambda(\theta)=|1-|\lambda_1||+|\lambda_1-\lambda_2|$, 其中

$$|\lambda_1-\lambda_2|=2\left|\sqrt{1-4r^2\sin^2\theta}\right|\Big/(1+2r).$$

当 λ_1,λ_2 是实根时, $|\lambda_1|=2r|\cos\theta|/(1+2r)+\left|\sqrt{1-4r^2\sin^2\theta}\right|\Big/(1+2r)$, 所以

$$\begin{aligned}\Lambda(\theta)&=1-2r|\cos\theta|/(1+2r)+\left|\sqrt{1-4r^2\sin^2\theta}\right|\Big/(1+2r)\geqslant 1-2r/(1+2r)\\ &=1/(1+2r)>0.\end{aligned}$$

当 λ_1, λ_2 是复根时, $\lambda_2 = \overline{\lambda}_1, |\lambda_1| = \sqrt{|\lambda_1\lambda_2|} = \sqrt{|(1-2r)/(1+2r)|}$, 从而

$$\Lambda(\theta) \geqslant 1 - |\lambda_1| = 1 - \sqrt{|1-2r|/|1+2r|} > 0.$$

可见函数 $\Lambda(\theta)$ 对 $\forall r > 0$ 于 $[0, 2\pi]$ 有正的下界 $m > 0$.

另一方面,

$$\boldsymbol{G}(\theta) - \frac{1}{2}(g_{11} + g_{22})\boldsymbol{I} = \begin{bmatrix} \dfrac{4r}{1+2r}\cos\theta & \dfrac{1-2r}{1+2r} \\ 1 & \dfrac{2r}{1+2r}\cos\theta \end{bmatrix}$$

其 F- 范数显然有上界 $K > 0$, 故条件 $(\beta)'$ 成立. 由定理 4.3 的推论, 可知 (3.4.8) 对 $\forall r > 0$ 都稳定.

判断差分格式稳定性最完整的代数准则已由 Kreiss 得到 (参看 [28]), 但这些准则难以检验, 所以有不少工作研究便于应用的充分条件 (参看 [14], [28]).

习 题

1. 证明实系数二次方程 $\lambda^2 - b\lambda - c = 0$ 的根按模小于或等于 1 的充要条件是

$$|b| \leqslant 1 - c \leqslant 2.$$

2. 证明差分格式 (3.1.15) 恒稳定.
3. 证明差分格式

$$\begin{cases} u_j^{n+1} - u_j^n = ar(u_{j+1}^n - u_j^n - u_j^{n+1} + u_{j-1}^{n+1}), \\ u_j^{n+2} - u_j^{n+1} = ar(u_{j+1}^{n+2} - u_j^{n+2} - u_j^{n+1} + u_{j-1}^{n+1})(a > 0) \end{cases}$$

(Saul'yev,1957) 恒稳定.

4. 考查如下隐 – 显格式:

$$\begin{cases} \dfrac{u_j^{2m-1} - u_j^{2m-2}}{\tau} = a\dfrac{u_{j+1}^{2m-2} - 2u_j^{2m-2} + u_{j-1}^{2m-2}}{h^2}, \\ \dfrac{u_j^{2m} - u_j^{2m-1}}{\tau} = a\dfrac{u_{j+1}^{2m} - 2u_j^{2m} + u_{j-1}^{2m}}{h^2}, \\ u_j^0 = \phi_j, \quad j = 1, 2, \cdots, J-1, \quad m = 1, 2, \cdots, M \\ u_0 = u_J = 0. \end{cases} \qquad (a > 0)$$

其计算量大约是向后差分格式的一半. 试证明它是恒稳定的.

*§5 变系数抛物方程

现在考虑平面有界域 G 上的抛物方程:

$$\frac{\partial u}{\partial t} = \nabla(a\nabla u) + f, \quad (x, y) \in G, \quad 0 < t \leqslant T, \tag{3.5.1}$$

其中 a, f 都是区域 G 上给定的光滑函数, $a=a(x,y)\geqslant\delta>0$. 初值条件为

$$u(x,y,0)=\varphi(x,y), \tag{3.5.2}$$

边值条件为下列三种类型之一:

$$u|_{\partial G}=\alpha(x,y), \quad (第一边值条件) \tag{3.5.3$_1$}$$

$$\left.\frac{\partial u}{\partial \boldsymbol{\nu}}\right|_{\partial G}=\beta(x,y), \quad (第二边值条件) \tag{3.5.3$_2$}$$

$$\left.\frac{\partial u}{\partial \boldsymbol{\nu}}+\mu u\right|_{\partial G}=\gamma(x,y), \quad (第三边值条件) \tag{3.5.3$_3$}$$

其中 $\boldsymbol{\nu}$ 是边界 ∂G 的单位外法向. 假定边值问题 (3.5.1) ~ $(3.5.3)_i$ $(i=1,2,3)$ 适定, 即解存在、唯一, 并且充分光滑.

像第二章 §3 那样, 我们用有限体积法构造差分格式. 取步长 h_1 和 $h_2, h=(h_1^2+h_2^2)^{1/2}$, 作两族与坐标轴平行的直线

$$x=ih_1, \quad i=0,\pm1,\cdots,$$
$$y=jh_2, \quad j=0,\pm1,\cdots,$$

将区域 G 剖分成一些小矩形和曲边三角形之并集. 以 G_h 表示属于 G 的内节点 $(x_i,y_j)=(ih_1,jh_2)$ 集合, Γ_h 表示网线 $x=x_i$ 或 $y=y_j$ 与边界 $\Gamma=\partial G$ 相交的界点集合, $\overline{G}_h=G_h\bigcup\Gamma_h$. 像第二章 §3 那样, 将 G_h 中的点分为正则内点和非正则内点 (参看第二章图 3.1). 记 $x_{i-\frac12}=\left(i-\frac12\right)h_1, y_{j-\frac12}=\left(j-\frac12\right)h_2$, 再作两族平行于坐标轴的直线 $x=x_{i-\frac12}$ 和 $y=y_{j-\frac12}$, 其交点属于 G 内部者为对偶剖分的内点, 直线与 Γ 的交点为对偶剖分的界点.

如第二章图 3.3, 对于任一正则内点 (x_i,y_j), 取对偶剖分的网点 $A(x_{i-\frac12}, y_{j-\frac12}), B(x_{i+\frac12},y_{j-\frac12}), C(x_{i+\frac12},y_{j+\frac12}), D(x_{i-\frac12},y_{j+\frac12})$. 用 G_{ij} 表示以 A,B,C,D 为顶点的矩形域, ∂G_{ij} 为其边界. 于 G_{ij} 积分抛物方程 (3.5.1) 两端, 并利用 Green 公式, 则得 (3.5.1) 的积分守恒形式:

$$\iint\limits_{G_{ij}}\frac{\partial u}{\partial t}\mathrm{d}x\mathrm{d}y=\int_{\partial G_{ij}}a\frac{\partial u}{\partial\boldsymbol{\nu}}\mathrm{d}s+\iint\limits_{G_{ij}}f\mathrm{d}x\mathrm{d}y,$$

式中 $\dfrac{\partial u}{\partial\boldsymbol{\nu}}$ 表示 u 沿 ∂G_{ij} 的外法向导数.

像第二章 §3 做过的那样, 用中矩形公式代替重积分和沿 ∂G_{ij} 四边的线积分, 用中心差商代替方向导数, 并除以 h_1h_2, 得

$$\begin{aligned}\frac{\partial u_{ij}}{\partial t}=&\frac{1}{h_1^2}\left[a_{i+\frac12,j}(u_{i+1,j}-u_{i,j})-a_{i-\frac12,j}(u_{i,j}-u_{i-1,j})\right]+\\&\frac{1}{h_2^2}\left[a_{i,j+\frac12}(u_{i,j+1}-u_{i,j})-a_{i,j-\frac12}(u_{i,j}-u_{i,j-1})\right]+f_{ij}.\end{aligned} \tag{3.5.4}$$

取时间步长 $\tau > 0$, 分别用向前和向后差商代替对 t 的偏导数, 则得向前差分格式:

$$u_{ij}^{n+1} = r_1 \left[a_{i+\frac{1}{2},j}(u_{i+1,j}^n - u_{i,j}^n) - a_{i-\frac{1}{2},j}(u_{i,j}^n - u_{i-1,j}^n)\right] + r_2 \left[a_{i,j+\frac{1}{2}}(u_{i,j+1}^n - u_{i,j}^n) - a_{i,j-\frac{1}{2}}(u_{i,j}^n - u_{i,j-1}^n)\right] + u_{i,j}^n + \tau f_{ij}, \tag{3.5.5}$$

向后差分格式:

$$u_{ij}^{n+1} = r_1 \left[a_{i+\frac{1}{2},j}(u_{i+1,j}^{n+1} - u_{i,j}^{n+1}) - a_{i-\frac{1}{2},j}(u_{i,j}^{n+1} - u_{i-1,j}^{n+1})\right] + r_2 \left[a_{i,j+\frac{1}{2}}(u_{i,j+1}^{n+1} - u_{i,j}^{n+1}) - a_{i,j-\frac{1}{2}}(u_{i,j}^{n+1} - u_{i,j-1}^{n+1})\right] + u_{i,j}^n + \tau f_{ij}, \tag{3.5.6}$$

其中 $r_1 = \tau/h_1^2, r_2 = \tau/h_2^2$. 它们的截断误差的阶都是 $O(\tau + h^2)$. 若将向前和向后差分格式作算术平均, 则得 Crank–Nicolson 格式, 其截断误差的阶为 $O(\tau^2 + h^2)$.

对于非正则内点 (x_i, y_j), 仍有类似 (3.5.4) 的方程. 例如设 (x_i, y_j) 是形如第二章图 3.4 所示的非正则内点 “×”, 则

$$\frac{\partial u_{ij}}{\partial t} = \frac{1}{\overline{h}_1}\left[a_{i+\frac{1}{2},j}\frac{u_{i+1,j} - u_{i,j}}{h_1} - \overline{a}_{i-\frac{1}{2},j}\frac{u_{i,j} - u_{i-1,j}}{h_1^-}\right] + \frac{1}{h_2^2}\left[a_{i,j+\frac{1}{2}}(u_{i,j+1} - u_{i,j}) - a_{i,j-\frac{1}{2}}(u_{i,j} - u_{i,j-1})\right] + f_{ij}, \tag{3.5.7}$$

其中 h_1^- 是 (x_i, y_j) 与界点 (x_{i-1}, y_j) 的距离, $\overline{a}_{i-\frac{1}{2},j}$ 等于 a 在 (x_i, y_j) 和 (x_{i-1}, y_j) 的一中间点的值, $\overline{h}_1 = \frac{1}{2}(h_1 + h_1^-)$.

(3.5.7) 相应的向前差分格式是

$$u_{ij}^{n+1} = \frac{\tau}{\overline{h}_1 h_1} a_{i+\frac{1}{2},j}(u_{i+1,j}^n - u_{i,j}^n) - \frac{\tau}{\overline{h}_1 h_1^-}\overline{a}_{i-\frac{1}{2},j}(u_{i,j}^n - u_{i-1,j}^n) + r_2 \left[a_{i,j+\frac{1}{2}}(u_{i,j+1}^n - u_{i,j}^n) - a_{i,j-\frac{1}{2}}(u_{i,j}^n - u_{i,j-1}^n)\right] + u_{i,j}^n + \tau f_{ij}, \tag{3.5.8}$$

向后差分格式是

$$u_{ij}^{n+1} = \frac{\tau}{\overline{h}_1 h_1} a_{i+\frac{1}{2},j}(u_{i+1,j}^{n+1} - u_{i,j}^{n+1}) - \frac{\tau}{\overline{h}_1 h_1^-}\overline{a}_{i-\frac{1}{2},j}(u_{i,j}^{n+1} - u_{i-1,j}^{n+1}) + r_2 \left[a_{i,j+\frac{1}{2}}(u_{i,j+1}^{n+1} - u_{i,j}^{n+1}) - a_{i,j-\frac{1}{2}}(u_{i,j}^{n+1} - u_{i,j-1}^{n+1})\right] + u_{i,j}^n + \tau f_{ij}. \tag{3.5.9}$$

在界点 $(x_i, y_j) \in \Gamma_h$, 若给的是第一边值条件, 则

$$u_{ij}^n = \alpha(x_i, y_j). \tag{3.5.10}$$

若给的是第二、三边值条件, 则仍按第二章 §3 的积分插值法在 (x_i, y_j) 建立一补充方程. 例如设 (x_i, y_j) 形如第二章图 3.5 中的 P_0, 边值条件为 $(3.5.3)_3$, 则差分

方程为 (参看第二章 (2.3.16))

$$(3.5.11)\quad \frac{u_{ij}^{n+1}-u_{ij}^{n}}{\tau}\widetilde{\Delta}ABC = \frac{\overline{AB}}{h_2}a_{i,j-\frac{1}{2}}(u_{i,j-1}^{n+l}-u_{ij}^{n+l}) + \frac{\overline{BC}}{h_1}a_{i+\frac{1}{2},j}(u_{i+1,j}^{n+l}-u_{ij}^{n+l}) + \widehat{CA}a_{ij}(\gamma_{ij}-\mu_{ij}u_{ij}^{n+l}) + \widetilde{\Delta}ABCf_{ij},\quad l=0,1.$$

利用正则点、非正则点和界点上的三类方程, 可组成各种计算方案. 考虑到非正则点附近的步长 $h_1^-,\overline{h}_1$ 一般小于 h_1, 故通常采用向后差分格式, 以免破坏正则点上的稳定条件. 现在给出几种实用的计算方案:

(一) 第一边值问题

方案 (I) 由 (3.5.5), (3.5.9) 和 (3.5.10) 组成的显格式. 先计算正则点, 再计算非正则点, 无需求解方程组.

方案 (II) 由 (3.5.6), (3.5.9) 和 (3.5.10) 组成的隐格式.

(二) 第二、三边值问题

方案 (III) 由 (3.5.5), (3.5.9) 和 (3.5.11) ($l=0$) 组成的显格式, 先计算正则内点, 再利用 (3.5.11) ($l=0$) 消去界点, 最后计算非正则点, 无需求解方程组.

方案 (IV) 由 (3.5.6), (3.5.9) 和 (3.5.11) ($l=1$) 组成的隐格式.

最后研究差分格式的稳定性和误差估计.

为简便计, 在方案 (I) 中用 (3.5.8) 代替 (3.5.9), 并假定 $h_1^-=\overline{h}_1=h_1$, 也就是说在正则和非正则点均用向前差分格式 (3.5.5). 设边值条件及右端项为齐次, 将 (3.5.5) 改写为

$$(3.5.12)\quad u_{ij}^{n+1} = [1-r_1(a_{i+\frac{1}{2},j}+a_{i-\frac{1}{2},j})-r_2(a_{i,j+\frac{1}{2}}+a_{i,j-\frac{1}{2}})]u_{i,j}^n + r_1a_{i+\frac{1}{2},j}u_{i+1,j}^n + r_1a_{i-\frac{1}{2},j}u_{i-1,j}^n + r_2a_{i,j+\frac{1}{2}}u_{i,j+1}^n + r_2a_{i,j-\frac{1}{2}}u_{i,j-1}^n.$$

定义最大模

$$\|\boldsymbol{U}\|_\infty = \max_{G_h}|u_{ij}|,$$

其中 $\boldsymbol{U}$ 是以 u_{ij} 为分量的向量. 设网比 r_1,r_2 满足条件

$$1-r_1(a_{i+\frac{1}{2},j}+a_{i-\frac{1}{2},j})-r_2(a_{i,j+\frac{1}{2}}+a_{i,j-\frac{1}{2}})\geqslant 0$$

或

$$(3.5.13)\quad r_1+r_2\leqslant 1/2\max_{G}a\,(x,y)\,,$$

则由 (3.5.12) 得

$$\|\boldsymbol{U}^{n+1}\|_\infty \leqslant \|\boldsymbol{U}^n\|_\infty \leqslant\cdots\leqslant \|\boldsymbol{U}^0\|_\infty,$$

足见 (3.5.12) 稳定.

以 $e_{ij}^n = u(x_i, y_j, t_n) - u_{ij}^n$ 表示差分解的误差, $R_{ij}^n(u)$ 表示差分格式的截断误差, 则 e_{ij}^n 满足

$$
\begin{aligned}
e_{ij}^{n+1} = &[1 - r_1(a_{i+\frac{1}{2},j} + a_{i-\frac{1}{2},j}) - r_2(a_{i,j+\frac{1}{2}} + a_{i,j-\frac{1}{2}})]e_{ij}^n + \\
&r_1 a_{i+\frac{1}{2},j} e_{i+1,j}^n + r_1 a_{i-\frac{1}{2},j} e_{i-1,j}^n + r_2 a_{i,j+\frac{1}{2}} e_{i,j+1}^n + r_2 a_{i,j-\frac{1}{2}} e_{i,j-1}^n + \tau R_{ij}^n.
\end{aligned}
$$

设稳定性条件 (3.5.13) 成立, 又 $|R_{ij}^n| \leqslant K(\tau + h^2)$, 则

$$
||\boldsymbol{E}^{n+1}||_\infty \leqslant ||\boldsymbol{E}^n||_\infty + K\tau(\tau + h^2),
$$

$\boldsymbol{E}^n = \left\{e_{ij}^n\right\}$ 是误差向量. 将上述不等式就 $n\,(n\tau \leqslant T)$ 相加, 得

$$
||\boldsymbol{E}^n||_\infty \leqslant ||\boldsymbol{E}^0||_\infty + KT(\tau + h^2), \quad 0 < n\tau \leqslant T. \tag{3.5.14}
$$

其次讨论方案 (Ⅱ)(隐格式) 的稳定性. 为简便计, 仍然假定 $h_1^- = \overline{h}_1 = h_1$, 即在正则和非正则点均用向后差分格式 (3.5.6). 仍设边值和右端项等于 0. 将 (3.5.6) 写成

$$
\begin{aligned}
&(1 + r_1 a_{i+\frac{1}{2},j} + r_1 a_{i-\frac{1}{2},j} + r_2 a_{i,j+\frac{1}{2}} + r_2 a_{i,j-\frac{1}{2}}) u_{ij}^{n+1} - \\
&r_1 a_{i+\frac{1}{2},j} u_{i+1,j}^{n+1} - r_1 a_{i-\frac{1}{2},j} u_{i-1,j}^{n+1} - r_2 a_{i,j+\frac{1}{2}} u_{i,j+1}^{n+1} - r_2 a_{i,j-\frac{1}{2}} u_{i,j-1}^{n+1} = u_{ij}^n.
\end{aligned} \tag{3.5.15}
$$

显然 (3.5.15) 满足第二章 §5 习题 2(该题的二维情形) 所述的条件, 因此

$$
||\boldsymbol{U}^{n+1}||_\infty \leqslant ||\boldsymbol{U}^n||_\infty \leqslant \cdots \leqslant ||\boldsymbol{U}^0||_\infty,
$$

所以 (3.5.15) 稳定.

若将上述方法用于与 (3.5.6) 相应的误差方程, 则得误差估计

$$
||\boldsymbol{E}^n||_\infty \leqslant ||\boldsymbol{E}^0||_\infty + KT(\tau + h^2).
$$

注 5.1 刚才我们用极值定理证明了向后差分格式恒稳定, 并且得到了误差估计. 但是有的差分格式, 比如 Crank Nicolson 格式, 不满足极值定理的条件, 这时论证稳定性和得到误差估计就困难了. 有兴趣的读者可参看 [12] 的第一章.

习 题

将第二章 §5 习题 2 的结果推广到二维情形.

§6 分数步长法

一维抛物方程的隐格式恒稳定, 用消元法求解的计算量也不大。对高维抛物方程, 隐格式虽然恒稳定, 但计算量增加很多; 显格式计算简单, 但稳定性条

件比一维情形苛刻, 于是人们企图构造一种恒稳定的格式, 使每层计算能分解为若干一维隐格式的形式. 本节介绍的各种**分数步长法**, 包括交替方向隐式法 —— **ADI 法** (Alternating direct implicit method)、**预 – 校法** (Predictor-corrector method) 和局部一维法 —— **LOD 法** (Locally one dimensional method) 就属于这一类.

6.1 ADI 法

作为模型, 考虑二维热传导方程的边值问题:

$$(3.6.1)\qquad \begin{cases} u_t = u_{xx} + u_{yy}, \quad 0 < x, \quad y < l, \quad t > 0, \\ u(x,y,0) = \varphi(x,y), \\ u(0,y,t) = u(l,y,t) = u(x,0,t) = u(x,l,t) = 0. \end{cases}$$

取空间步长 $h = l/M$, 时间步长 $\tau > 0$, 作两族平行于坐标轴的网线: $x = x_j = jh, y = y_k = kh, j,k = 0,1,\cdots,M$, 将区域 $0 \leqslant x, y \leqslant l$ 分割成 M^2 个小矩形. 第一个 **ADI 法(交替方向隐格式)** 是 Peaceman 和 Rachford(1955) 提出的, 他们把由第 n 层到第 $n+1$ 层计算分成两步: 先由第 n 层到第 $n+\frac{1}{2}$ 层, 对 u_{xx} 用向后差分逼近, 对 u_{yy} 用向前差分逼近, 然后由第 $n+\frac{1}{2}$ 到第 $n+1$ 层, 对 u_{xx} 用向前差分逼近, 对 u_{yy} 用向后差分逼近, 于是得到如下 ADI 格式 (PR 格式):

$$(3.6.2)_1\qquad \frac{u_{jk}^{n+\frac{1}{2}} - u_{jk}^n}{\tau/2} = \frac{1}{h^2}\left(\delta_x^2 u_{jk}^{n+\frac{1}{2}} + \delta_y^2 u_{jk}^n\right),$$

$$(3.6.2)_2\qquad \frac{u_{jk}^{n+1} - u_{jk}^{n+\frac{1}{2}}}{\tau/2} = \frac{1}{h^2}\left(\delta_x^2 u_{jk}^{n+\frac{1}{2}} + \delta_y^2 u_{jk}^{n+1}\right),$$

其中 $j,k = 1,2,\cdots,M-1, n = 0,1,2,\cdots$, 上标 $n+\frac{1}{2}$ 表示在 $t = t_{n+\frac{1}{2}} = \left(n+\frac{1}{2}\right)\tau$ 取值. 假定第 n 层的 u_{jk}^n 已求得, 则由 $(3.6.2)_1$ 求出 $u_{jk}^{n+\frac{1}{2}}$, 这只需按行 $(j = 1,2,\cdots,M-1)$ 解一些具三对角系数矩阵的方程组; 再由 $(3.6.2)_2$ 求出 u_{jk}^{n+1}, 这只需按列 $(k = 1,2,\cdots,M-1)$ 解一些具三对角系数矩阵的方程组, 所以计算是容易实现的.

现在估计截断误差的阶. 将 $(3.6.2)_1, (3.6.2)_2$ 相加、相减, 依次得

$$(3.6.3)_1\qquad \frac{u_{jk}^{n+1} - u_{jk}^n}{\tau/2} = \frac{2}{h^2}\delta_x^2 u_{jk}^{n+\frac{1}{2}} + \frac{1}{h^2}\delta_y^2(u_{jk}^n + u_{jk}^{n+1}),$$

和

$$(3.6.3)_2\qquad 4u_{jk}^{n+\frac{1}{2}} = 2(u_{jk}^{n+1} + u_{jk}^n) - \frac{\tau}{h^2}\delta_y^2(u_{jk}^{n+1} - u_{jk}^n).$$

消去过渡值 $u_{jk}^{n+\frac{1}{2}}$, 则

$$\left(I+\frac{1}{4}\frac{\tau^2}{h^4}\delta_x^2\delta_y^2\right)\frac{u_{jk}^{n+1}-u_{jk}^n}{\tau}=\frac{1}{h^2}(\delta_x^2+\delta_y^2)\frac{u_{jk}^{n+1}+u_{jk}^n}{2}. \tag{3.6.4}$$

以 $\overline{u}_{jk}^n$ 表示精确解在节点 (x_j,y_k,t_n) 的值 $u(x_j,y_k,t_n)$, 利用 Taylor 展式, 易见

$$\left(I+\frac{1}{4}\frac{\tau^2}{h^4}\delta_x^2\delta_y^2\right)\frac{\overline{u}_{jk}^{n+1}-\overline{u}_{jk}^n}{\tau}=\frac{1}{h^2}(\delta_x^2+\delta_y^2)\frac{\overline{u}_{jk}^{n+1}+\overline{u}_{jk}^n}{2}+O(\tau^2+h^2). \tag{3.6.5}$$

与 (3.6.4) 比较, 可见截断误差的阶为 $O(\tau^2+h^2)$.

为了检验格式 $(3.6.2)_{1-2}$ 的稳定性, 我们将它改写成

$$\left(I-\frac{r}{2}\delta_x^2\right)u_{jk}^{n+\frac{1}{2}}=\left(I+\frac{r}{2}\delta_y^2\right)u_{jk}^n \quad \left(r=\frac{\tau}{h^2}\right), \tag{$(3.6.2)_1'$}$$

$$\left(I-\frac{r}{2}\delta_y^2\right)u_{jk}^{n+1}=\left(I+\frac{r}{2}\delta_x^2\right)u_{jk}^{n+\frac{1}{2}} \quad \left(r=\frac{\tau}{h^2}\right). \tag{$(3.6.2)_2'$}$$

注意, $-\delta_x^2$ 和 $-\delta_y^2$ 都是正定矩阵, 故以上二式左端的差分算子有逆. 消去过渡层, 即得

$$\left(I-\frac{r}{2}\delta_x^2\right)\left(I-\frac{r}{2}\delta_y^2\right)u_{jk}^{n+1}=\left(I+\frac{r}{2}\delta_x^2\right)\left(I+\frac{r}{2}\delta_y^2\right)u_{jk}^n. \tag{3.6.6}$$

这是 (3.6.4) 的一个等价形式. 以 $u_{jk}^n=v^n\exp\mathrm{i}(\alpha x_j+\beta y_k)$ 代到 (3.6.5) 或 $(3.6.2)_{1-2}'$, 消去公因子, 则知增长因子

$$G(\alpha h,\beta h)=\frac{\left(I-2r\sin^2\frac{\alpha h}{2}\right)\left(I-2r\sin^2\frac{\beta h}{2}\right)}{\left(I+2r\sin^2\frac{\alpha h}{2}\right)\left(I+2r\sin^2\frac{\beta h}{2}\right)}.$$

显然对任何 $r>0, |G|\leqslant 1$, 故 ADI 法绝对稳定.

总之, 在计算量、截断误差的阶和稳定性方面,ADI 法都是令人满意的, 因此受到人们的普遍注意.

注 6.1 若边值问题 (3.6.1) 的边值为非齐, 比如在边界上 $u=\alpha(x,y,t)$. 则过渡层 $u_{ij}^{n+\frac{1}{2}}$ 的边值取为

$$u_{ij}^{n+\frac{1}{2}}=\frac{1}{2}(\alpha(x_i,y_j,t_n)+\alpha(x_i,y_j,t_{n+1}))$$

或按 $(3.6.3)_2$ 取为

$$u_{ij}^{n+\frac{1}{2}}=\frac{1}{2}(\alpha(x_i,y_j,t_n)+\alpha(x_i,y_j,t_{n+1}))-\frac{r}{4}\delta_y^2(\alpha(x_i,y_j,t_{n+1})-\alpha(x_i,y_j,t_n)).$$

注 6.2 若边值问题 (3.6.1) 的右端非齐, 比如等于 $f(x,y,t)$, 则 (3.6.4) 右端增加一项 $f_{jk}^{n+\frac{1}{2}}=f(x_j,y_k,t_{n+\frac{1}{2}})$, 相应的 ADI 格式 $(3.6.2)_1'$, $(3.6.2)_2'$ 右端分别加一项 $\frac{1}{2}f_{jk}^{n+\frac{1}{2}}$.

注 6.1 和注 6.2 也可用于以后介绍的预 – 校法和分数步长法.

6.2 预 – 校法

考虑逼近 (3.6.1) 的向后差分格式:

$$\frac{u_{jk}^{n+1}-u_{jk}^{n}}{\tau}=\frac{1}{h^2}(\delta_x^2 u_{jk}^{n+1}+\delta_y^2 u_{jk}^{n+1}).$$

令 $r=\tau/h^2$, 将它改写成

$$(I-r\delta_x^2-r\delta_y^2)u_{jk}^{n+1}=u_{jk}^{n}.$$

两端各加一项 $r^2\delta_x^2\delta_y^2u_{jk}^{n+1}$, 得

$$(I-r\delta_x^2-r\delta_y^2+r^2\delta_x^2\delta_y^2)u_{jk}^{n+1}=r^2\delta_x^2\delta_y^2u_{jk}^{n+1}+u_{jk}^{n},$$

再略去右端高阶无穷小项:

$$r^2\delta_x^2\delta_y^2u_{jk}^{n+1}=h^4r^2\left(\frac{\delta_x^2}{h^2}\right)\left(\frac{\delta_y^2}{h^2}\right)u_{jk}^{n+1},$$

并将左端分解, 则得

$$(I-r\delta_x^2)(I-r\delta_y^2)u_{jk}^{n+1}=u_{jk}^{n}. \tag{3.6.7}$$

这是恒稳定的格式, 它逼近 (3.6.1) 的截断误差的阶为 $O(\tau+h^2)$.

将 (3.6.7) 用于区间 $[t_n,t_{n+\frac12}]$, 得

$$\left(I-\frac{\tau}{2h^2}\delta_x^2\right)\left(I-\frac{\tau}{2h^2}\delta_y^2\right)u_{jk}^{n+\frac12}=u_{jk}^{n}.$$

引进 $u_{jk}^{n+\frac14}=\left(I-\frac{\tau}{2h^2}\delta_y^2\right)u_{jk}^{n+\frac12}$, 则得计算 $u_{jk}^{n+\frac12}$ 的预算格式:

$$\left(I-\frac{\tau}{2h^2}\delta_x^2\right)u_{jk}^{n+\frac14}=u_{jk}^{n}, \tag{3.6.8$_1$}$$

$$\left(I-\frac{\tau}{2h^2}\delta_y^2\right)u_{jk}^{n+\frac12}=u_{jk}^{n+\frac14}. \tag{3.6.8$_2$}$$

可将 (3.6.8) 写成更直观的形式:

$$\frac{u_{jk}^{n+\frac14}-u_{jk}^{n}}{\tau/2}=\frac{1}{h^2}\delta_x^2u_{jk}^{n+\frac14}, \tag{3.6.9$_1$}$$

$$\frac{u_{jk}^{n+\frac12}-u_{jk}^{n+\frac14}}{\tau/2}=\frac{1}{h^2}\delta_y^2u_{jk}^{n+\frac12}. \tag{3.6.9$_2$}$$

其次, 利用 $u_{jk}^{n},u_{jk}^{n+\frac12}$ 构造更精确的校正格式:

$$\frac{u_{jk}^{n+1}-u_{jk}^{n}}{\tau}=\frac{1}{h^2}(\delta_x^2+\delta_y^2)u_{jk}^{n+\frac12}, \tag{3.6.10}$$

由此得到第 $n+1$ 层上的 u_{jk}^{n+1}.

预 – 校格式 $(3.6.9)_{1-2}$ 和 (3.6.10) 是由 Yanenko 建立的. 由 $(3.6.9)_{1-2}$ 消去 $u_{jk}^{n+\frac{1}{4}}$, 得

$$\left(I-\frac{r}{2}\delta_x^2\right)\left(I-\frac{r}{2}\delta_y^2\right)u_{jk}^{n+\frac{1}{2}}=u_{jk}^n.$$

再用上式消去 (3.6.10) 中的 $u_{jk}^{n+\frac{1}{2}}$, 又得

$$u_{jk}^{n+1}-u_{jk}^n=r(\delta_x^2+\delta_y^2)\left(I-\frac{r}{2}\delta_y^2\right)^{-1}\left(I-\frac{r}{2}\delta_x^2\right)^{-1}u_{jk}^n.$$

注意 δ_x^2,δ_y^2 的矩阵形式相同, 故上式右端的差分算子可换序, 因此可化为

$$\left(I-\frac{r}{2}\delta_x^2\right)\left(I-\frac{r}{2}\delta_y^2\right)(u_{jk}^{n+1}-u_{jk}^n)=r(\delta_x^2+\delta_y^2)u_{jk}^n.$$

这就是 (3.6.6) 的等价形式. 由此可见, 预 – 校格式恒稳定, 截断误差的阶为 $O(\tau^2+h^2)$.

6.3 LOD 法

如果用 Crank–Nicolson 格式直接代替预–校法的 "向后差分格式", 则得到更简单的所谓 **LOD 格式 (局部一维格式)**:

$$(3.6.11)_1\qquad \frac{u_{jk}^{n+\frac{1}{2}}-u_{jk}^n}{\tau}=\frac{1}{2h^2}\delta_x^2(u_{jk}^{n+\frac{1}{2}}+u_{jk}^n),$$

$$(3.6.11)_2\qquad \frac{u_{jk}^{n+1}-u_{jk}^{n+\frac{1}{2}}}{\tau}=\frac{1}{2h^2}\delta_y^2(u_{jk}^{n+1}+u_{jk}^{n+\frac{1}{2}}).$$

消去过渡层 (注意 δ_x^2 和 δ_y^2 可换序), 又得到 (3.6.6) 的等价形式. 可见 LOD 格式 $(3.6.11)_{1-2}$ 恒稳定, 截断误差的阶是 $O(\tau^2+h^2)$.

$(3.6.11)_{1-2}$ 可写成更对称的所谓**双循环局部一维格式**:

$$(3.6.12)\qquad \begin{cases}\dfrac{u_{jk}^{n-\frac{1}{2}}-u_{jk}^{n-1}}{\tau}=\dfrac{1}{2h^2}\delta_x^2(u_{jk}^{n-\frac{1}{2}}+u_{jk}^{n-1}),\\[2mm] \dfrac{u_{jk}^{n}-u_{jk}^{n-\frac{1}{2}}}{\tau}=\dfrac{1}{2h^2}\delta_y^2(u_{jk}^{n}+u_{jk}^{n-\frac{1}{2}}),\\[2mm] \dfrac{u_{jk}^{n+\frac{1}{2}}-u_{jk}^{n}}{\tau}=\dfrac{1}{2h^2}\delta_y^2(u_{jk}^{n+\frac{1}{2}}+u_{jk}^{n}),\\[2mm] \dfrac{u_{jk}^{n+1}-u_{jk}^{n+\frac{1}{2}}}{\tau}=\dfrac{1}{2h^2}\delta_x^2(u_{jk}^{n+1}+u_{jk}^{n+\frac{1}{2}}).\end{cases}$$

习　　题

(实习题) 用 PR 格式的非齐形式求解:

$$\begin{cases} u_t = \nu(u_{xx} + u_{yy}) + f(x,t), (x,y) \in (0,1) \times (0,1), t > 0, \\ u(0,y,t) = u(1,y,t) = 0, y \in [0,1], t \geqslant 0, \\ u(x,0,t) = u(x,1,t) = 0, x \in [0,1], t \geqslant 0, \\ u(x,y,0) = 0, (x,y) \in [0,1] \times [0,1], \end{cases}$$

其中 $f(x,t) = \sin 5\pi t \sin 2\pi x \sin \pi y, \nu = 1$. 网格步长 $h_1 = h_2 = 0.1, 0.05, \tau = 0.01$. 计算 u 于 $t = 0.1, 0.2, 0.4, 0.8$ 的近似值.

§7 数值例子

7.1 一维抛物方程的初边值问题

求解一维抛物方程的初边值问题:

$$(3.7.1)_1 \qquad \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \sin t, \quad 0 < x < 1, \quad t > 0,$$

$$(3.7.1)_2 \qquad u_x(0,t) = u_x(1,t) = 0, \quad t > 0,$$

$$(3.7.1)_3 \qquad u(x,0) = \cos \pi x, \quad 0 < x < 1$$

(精确解 $u = \mathrm{e}^{-\pi^2 t} \cos \pi x + (1 - \cos t)$)

设空间步长 $h = 1/J$, 时间步长 $\tau > 0, t_n = n\tau$, 网比 $r = \tau/h^2$.

(一) 向前格式

$$\frac{u_j^{n+1} - u_j^n}{\tau} = \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{h^2} + \sin t_n, j = 1, \cdots, J-1, n = 1, \cdots, N.$$

边值条件:

$$j = 0, \frac{u_0^{n+1} - u_0^n}{\tau} = \frac{u_1^n - 2u_0^n + u_{-1}^n}{h^2} + \sin t_n, \quad u_{-1}^n = u_1^n,$$

$$j = J, \frac{u_J^{n+1} - u_J^n}{\tau} = \frac{u_{J+1}^n - 2u_J^n + u_{J-1}^n}{h^2} + \sin t_n, \quad u_{J+1}^n = u_{J-1}^n$$

初值条件

$$u(x_j, 0) = \cos \pi x_j, j = 0, 1, \cdots, J.$$

计算方案:

a) 取 $h = 1/40, \tau = 1/3\,200$, 此时 $r = 1/2$, 计算到时间层 $t_{3\,200} = 1$.

b) 取 $h = 1/80, \tau = 1/12\,800$, 此时 $r = 1/2$, 计算到时间层 $t_{12\,800} = 1$.
计算结果列于表 7.1, 其中 $x_j = j/4, j = 1, \cdots, 4$. 从表中看出, 计算是稳定的, 解随网格加细逼近精确解.

c) 取 $h = 1/80, \tau = 1/3\,200$, 此时 $r = 2$, 计算到时间层 $t_{20} = 0.006\,25, t_{40} = 0.012\,5$ 即发生激烈震荡, 结果不稳定.

(二) 向后格式

$$\frac{u_j^{n+1} - u_j^n}{\tau} = \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{h^2} + \sin t_{n+1}, j = 1, \cdots, J-1, n = 1, \cdots, N.$$

边值条件:

$$j = 0, \frac{u_0^{n+1} - u_0^n}{\tau} = \frac{u_1^{n+1} - 2u_0^{n+1} + u_{-1}^{n+1}}{h^2} + \sin t_{n+1}, \quad u_{-1}^{n+1} = u_1^{n+1},$$

$$j = J, \frac{u_J^{n+1} - u_J^n}{\tau} = \frac{u_{J+1}^{n+1} - 2u_J^{n+1} + u_{J-1}^{n+1}}{h^2} + \sin t_{n+1}, \quad u_{J+1}^{n+1} = u_{J-1}^{n+1}.$$

初值条件:

$$u(x_j, 0) = \cos \pi x_j, \quad j = 0, 1, \cdots, J.$$

计算方案:

a) $h = 1/40, \tau = 1/1\,600$, 此时 $r = 1$, 计算到时间层 $t_{1\,600} = 1$,

b) $h = 1/80, \tau = 1/3\,200$, 此时 $r = 2$, 计算到时间层 $t_{3\,200} = 1$.
计算结果列于表 7.1, 其中 $x_j = j/4, j = 1, \cdots, 4$. 从表中看出, 对大 r, 计算仍稳定, 但精度随网格变细增加缓慢.

表 7.1

格式	h	r	x_1	x_2	x_3	x_4	误差阶
向前	1/4	1/2	0.459 617	0.459 602	0.459 566	0.459 530	1.0×10^{-3}
	1/8	1/2	0.459 716	0.459 701	0.459 665	0.459 626	1.0×10^{-4}
向后	1/40	1	0.460 014	0.459 999	0.459 961	0.459 923	1.0×10^{-3}
	1/80	2	0.459 881	0.459 866	0.459 829	0.459 792	1.0×10^{-3}
对称	1/40	1	0.459 750	0.459 734	0.459 696	0.459 661	1.0×10^{-6}
	1/80	2	0.459 750	0.459 734	0.459 698	0.459 661	1.0×10^{-7}
精确解			0.459 749	0.459 734	0.459 698	0.459 661	

(三) 六点对称格式

$$\frac{u_j^{n+1} - u_j^n}{\tau} = \frac{1}{2h^2}\left(\left(u_{j+1}^{n+1} + u_{j+1}^n\right) - 2\left(u_j^{n+1} + u_j^n\right) + \left(u_{j-1}^{n+1} + u_{j-1}^n\right)\right) + \frac{1}{2}\left(\sin t_{n+1} + \sin t_n\right), \quad j = 1, \cdots, J-1, n = 1, \cdots, N.$$

边值条件: 在 $j=0,J$ 列 Crank-Nicolson 格式, 其中

$$u_{-1}^n+u_{-1}^{n+1}=u_1^n+u_1^{n+1},\quad u_{J+1}^n+u_{J+1}^{n+1}=u_{J-1}^n+u_{J-1}^{n+1}.$$

初值条件:

$$u(x_j,0)=\cos\pi x_j,\quad j=0,1,\cdots,J.$$

计算方案

a) $h=1/40,\tau=1/1\,600$, 此时 $r=1$, 计算到时间层 $t_{1\,600}=1$,

b) $h=1/80,\tau=1/3\,200$, 此时 $r=2$, 计算到时间层 $t_{3\,200}=1$.

实际按 15 位小数计算, 这里仅取 6 位列于表 7.1, 其中 $x_j=j/4,j=1,\cdots,4$ 从表中看出, 对大 r, 计算仍稳定, 且精度比向前格式和向后格式明显提高.

7.2 二维抛物方程的初边值问题

用六点对称格式, ADI 法, 预较法和 LOD 法求解二维抛物方程的的初边值问题:

$$(3.7.2)_1\qquad \frac{\partial u}{\partial t}=4^{-2}(u_{xx}+u_{yy}),(x,y)\in G=(0,1)\times(0,1),\quad t>0,$$

$$(3.7.2)_2\qquad u(0,y,t)=u(1,y,t)=0,\quad 0<y<1,\quad t>0,$$

$$(3.7.2)_3\qquad u_y(x,0,t)=u_y(x,1,t)=0,\quad 0<x<1,\quad t>0,$$

$$(3.7.2)_4\qquad u(x,y,0)=\sin\pi x\cos\pi y.$$

(精确解: $u=\sin\pi x\cos\pi y\exp\left(-\frac{\pi^2}{8}t\right)$)

设 $x_j=jh\,(j=0,1,\cdots,J),y_k=kh\,(k=0,1,\cdots,K),t_n=n\tau(n=0,1,\cdots,N)$, 差分解为 u_{jk}^n, 则边值条件为

$$u_{0k}^n=u_{Jk}^n=0,k=0,1,\cdots,K,$$
$$u_{j0}^n=u_{j1}^n,u_{j,K-1}^n=u_{j,K}^n,j=0,1,\cdots,J.$$

初值条件为

$$u_{jk}^0=\cos\pi x_j\sin\pi y_k.$$

取空间步长 $h=h_1=h_2=1/40$, 时间步长 $\tau=1/1\,600$, 网比 $r=\tau/h^2=1$. 用六点对称格式, ADI 法, 预校法和 LOD 法分别计算到时间层 $t=1$. 表 7.2 列出在节点

$$(x_j,y_k)=\left(\frac{j}{4},\frac{k}{4}\right),j,k=1,2,3.$$

的计算结果. 从表中看出, 四种方法的精度基本相同. 但三种分数步长法的计算量和存储量比 Crank-Nicolson 格式小许多. 图 7.1 是精确解曲面, 图 7.2 是 Crank-Nicolson 格式解曲面.

表7.2

方法	(x_j, y_k)	x_1	x_2	x_3	误差阶
六点对称格式	y_1	0.142 058	0.200 90	0.142 058	1.0×10^{-2}
ADI		0.142 059	0.200 902	0.142 059	1.0×10^{-2}
预较法		0.142 058	0.200 900	0.142 606	1.0×10^{-2}
LOD		0.142 058	0.200 900	0.142 058	1.0×10^{-2}
精确解		0.145 606	0.205 919	0.145 606	
六点对称格式	y_2	0.000 000	0.000 000	0.000 000	1.0×10^{-1}
ADI		0.000 000	0.000 000	0.000 000	1.0×10^{-1}
预较法		0.000 000	0.000 000	0.000 000	1.0×10^{-1}
LOD		0.000 000	0.000 000	0.000 000	1.0×10^{-1}
精确解		0.000 000	0.000 000	0.000 000	
六点对称格式	y_3	-0.142 058	-0.200 900	-0.142 058	1.0×10^{-2}
ADI		-0.142 059	-0.200 902	-0.142 059	1.0×10^{-2}
预较法		-0.142 058	-0.200 900	-0.142 606	1.0×10^{-2}
LOD		-0.142 058	-0.200 900	-0.142 058	1.0×10^{-2}
精确解		-0.145 606	-0.205 919	-0.145 606	

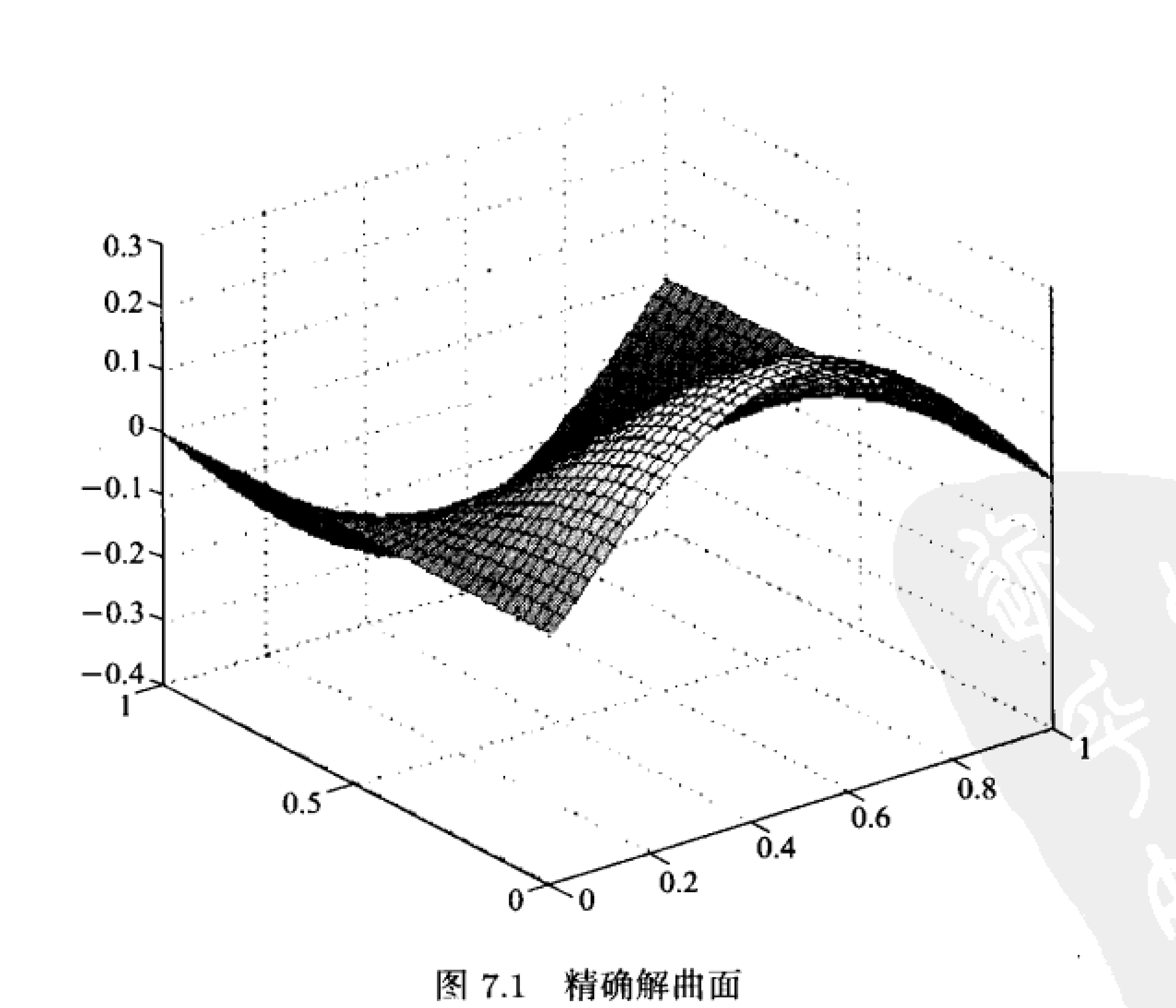


图 7.1　精确解曲面

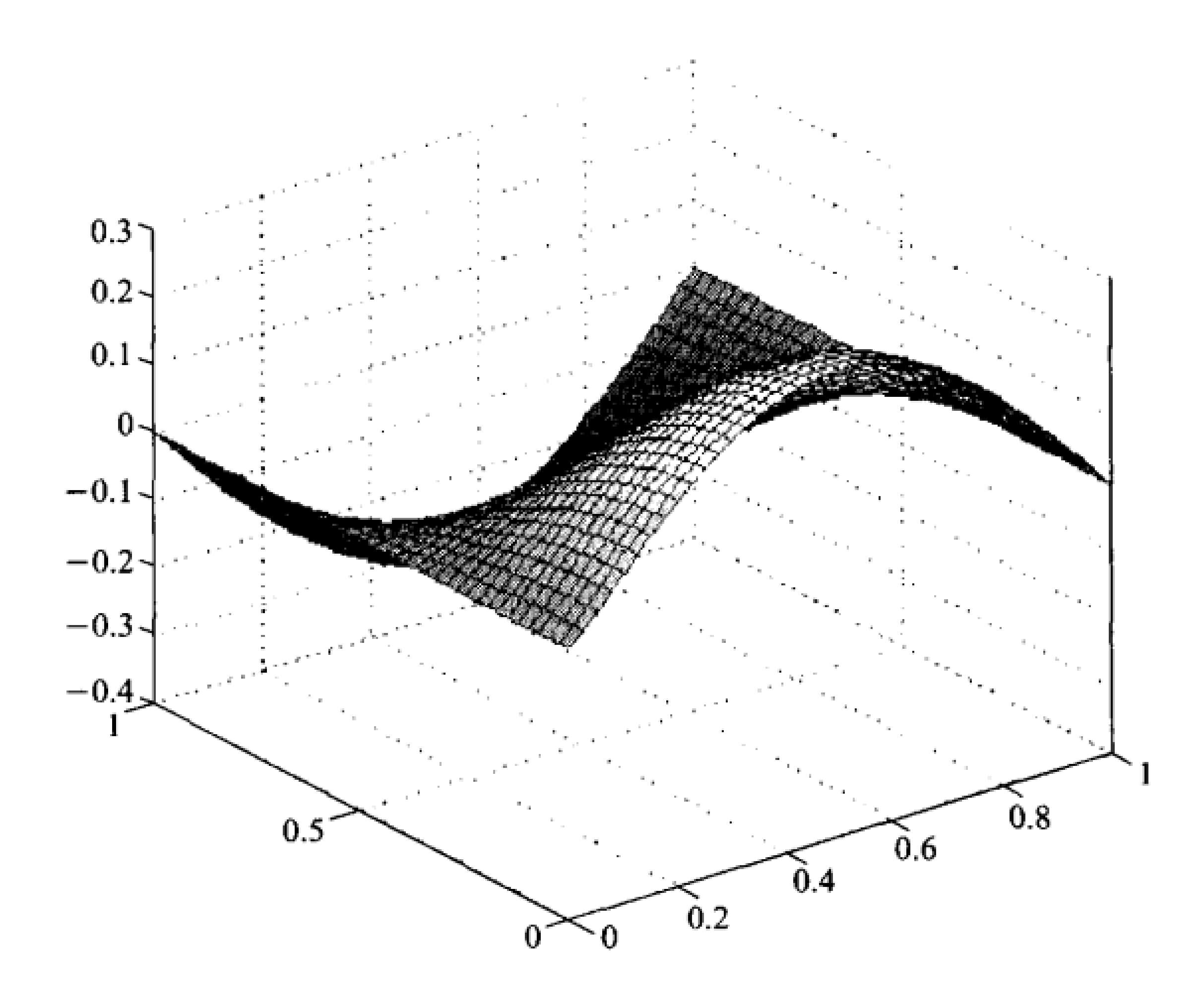


图 7.2 六点对称格式解曲面

7.3 含对流项的抛物方程

考虑含对流项的抛物方程:

$$\frac{\partial u}{\partial t}+b\frac{\partial u}{\partial x}=a\frac{\partial^2 u}{\partial x^2}, \tag{3.7.3}$$

其中 $a>0,b>0$ 是常数. 考虑中心差分格式

$$\frac{u_j^n-u_j^{n-1}}{\tau}+b\frac{u_{j+1}^{n-1}-u_{j-1}^{n-1}}{h}=a\frac{u_{j+1}^{n-1}-2u_j^{n-1}+u_{j-1}^{n-1}}{h^2}. \tag{3.7.4}$$

用 Fourier 方法, 知增长因子

$$\rho(\theta)=(1-2r_1)+2r_1\cos\theta-ir\sin\theta$$

其中

$$r=\frac{a\tau}{h^2},r_1=\frac{b\tau}{h}=\frac{a\tau}{h^2}\frac{bh}{a}=\frac{bh}{a}r.$$

经计算, 得

$$|\rho(\theta)|^2=(1-2r)^2+r_1^2+4r(1-2r)\cos\theta+\left(4r^2-r_1^2\right)\cos^2\theta,\quad 0\leqslant\theta\leqslant 2\pi. \tag{3.7.5}$$

令 $c=b/a$, 则

$$r_1^2=c^2r^2h^2=c^2ar\tau.$$

由定理 4.1, 为使格式稳定, 可略去 (3.7.5) 中的 $r_1^2 = O(\tau)$, 要求

$$(1-2r)^2 + 4r(1-2r)\cos\theta + 4r^2\cos^2\theta \leqslant 1,$$

即

$$(1-2r+2r\cos\theta)^2 \leqslant 1$$

或

$$-1 \leqslant 1 - 4r\sin^2\frac{1}{2}\theta \leqslant 1,$$

从而得稳定性条件为

$$r \leqslant \frac{1}{2}. \tag{3.7.6}$$

这是对充分小 h 得到的稳定性条件.

实际计算时步长总是有限的, 为得出相应的稳定性条件, 应要求

$$|\rho(\theta)|^2 \leqslant 1.$$

于 (3.7.5) 令 $x = \cos\theta$. 考虑抛物线

$$y = |\rho(\theta)|^2 = (1-2r)^2 + r_1^2 + 4r(1-2r)x + \left(4r^2 - r_1^2\right)x^2, \quad -1 \leqslant x \leqslant 1$$

它在区间 $[-1,1]$ 非负, 在端点取值:

$$y = \begin{cases} y_{-1} = (1-4r)^2, & \text{当 } x = -1, \\ y_1 = 1, & \text{当 } x = 1. \end{cases}$$

设 $(4r^2 - r_1^2) \geqslant 0$ 或 $\dfrac{r_1}{2} \leqslant r$, 则抛物线下凸, 所以 y 在端点取极大. 故稳定性要求 $y_{\pm1} \leqslant 1$ 即 $r \leqslant 1/2$, 从而

$$\left|\frac{r_1}{2}\right| \leqslant r \leqslant \frac{1}{2}. \tag{3.7.7}$$

其次设

$$4r^2 - r_1^2 < 0. \tag{3.7.8}$$

此时抛物性上凸, 极值点位于抛物线顶点

$$x_0 = -\frac{2r(1-2r)}{(4r^2 - r_1^2)},$$

极值为

$$y_0 = \frac{r_1^2}{4r^2 - r_1^2}\left(4r - r_1^2 - 1\right).$$

由 (3.7.7) 知 $x_0\,(\leqslant 0)$ 在 $x=1$ 左侧, 又 $y_{-1}\leqslant y_1$, 所以 y_0 是最大值. 为使

$$y_0=\frac{r_1^2}{4r^2-r_1^2}\left(4r-r_1^2-1\right)\leqslant 1$$

必须

$$(r_1-2r)^2\leqslant 0$$

从而 $r_1=2r$, 与假设 (3.7.8) 矛盾, 此时格式不稳定. 总之, 差分格式稳定的充要条件是 (3.7.7) 成立.

注 条件 (3.7.7) 的第一个不等式为

$$|b|\,h\leqslant 2a,$$

对充分小的 h 总成立. 对有限的 $h>0$, 若上述不等式不成立, 则解可能出现不应有的震荡, 但并不影响解收敛.

例如求解对流 (占优) 扩散方程

$$(3.7.9)_1 \qquad u_t+bu_x=au_{xx}\quad t>0,\quad 0<x<\infty,$$

其中 $a,b>0$ 是常数. 定解条件为

$$(3.7.9)_2 \qquad u\,(x,0)=0,\quad x>0,$$

$$(3.7.9)_3 \qquad u\,(0,t)=u_0,\quad t\geqslant 0,\qquad u_0:\ \text{正常数},$$

$$(3.7.9)_4 \qquad u\,(\infty,t)=0,\quad t\geqslant 0.$$

精确解为:

$$u\,(x,t)=\frac{u_0}{2}\left\{\operatorname{erfc}\left(\frac{x-bt}{2\sqrt{at}}\right)+\exp\left(\frac{bx}{a}\right)\operatorname{erfc}\left(\frac{x+bt}{2\sqrt{at}}\right)\right\},$$
$$\operatorname{erfc}(x)=\frac{2}{\sqrt{\pi}}\int_x^{\infty}\mathrm{e}^{-t^2}\mathrm{d}t.$$

取 $a=1$, 用显中心差分格式

$$\frac{u_j^{k+1}-u_j^k}{\tau}+b\frac{u_{j+1}^k-u_{j-1}^k}{2h}=\frac{u_{j+1}^k-2u_j^k+u_{j-1}^k}{h^2}$$

求解 (3.7.9).

(i) 设对流项 $b=1$

取步长 $h=0.5,\tau=0.125$, 此时网比 $r=0.5$ 满足稳定性条件 (3.7.6). 计算到前 50 层仍未发生震荡, 解曲线如图 7.3, 实黑线为精确解, 虚线为近似解.

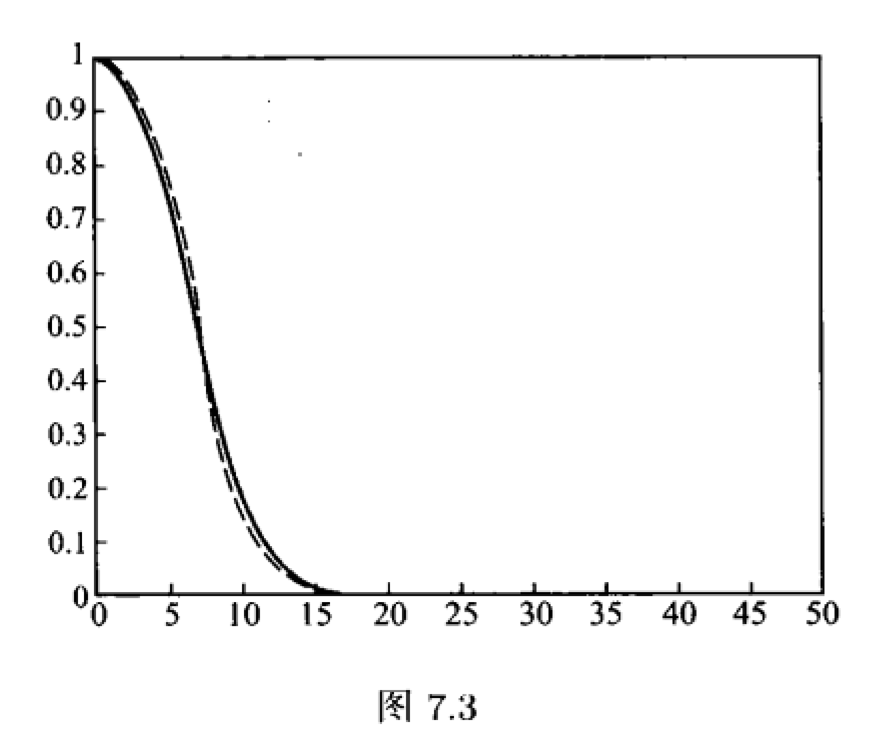


图 7.3

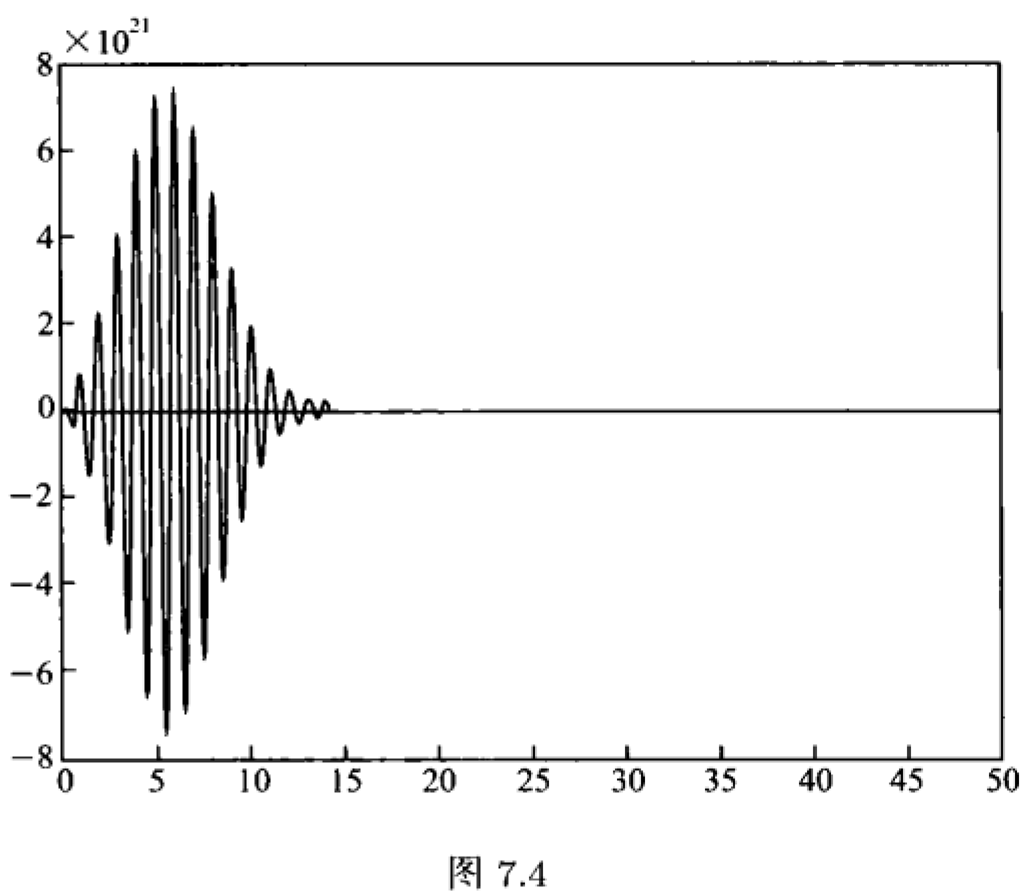


图 7.4

当取 τ=0.25 时, 网比 $r=1$ 不满足稳定条件 (2.7.6), 计算到第 50 层即发生震荡, 解曲线如图 7.4.

(ii) 当对流项 $b=10$ 时

取 $h=0.5, \tau=0.125$, 此时网比 $r=0.5$ 仍满足 (3.7.6), 但不满足 (3.7.7), 计算到第三层即发生震荡, 如图 7.5 中的虚线. 若步长进一步缩小, 只要 $r \leqslant 1/2$, 则差分解仍将收敛.

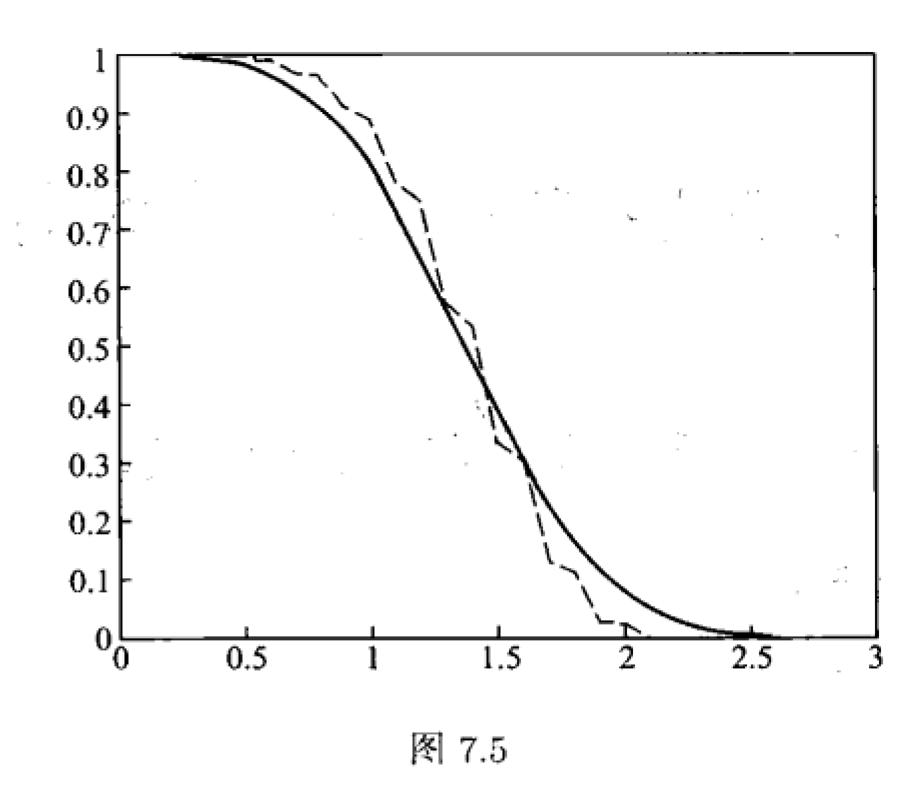


图 7.5

第四章　双曲型方程的有限差分法

§1　波动方程的差分逼近

1.1　波动方程及其特征

二阶线性双曲型偏微分方程的最简单模型是波动方程:

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \tag{4.1.1}$$

其中 $a > 0$ 是常数. 根据二阶偏微分方程理论, 与 (4.1.1) 相应的特征方程为

$$\mathrm{d}x^2 - a^2 \mathrm{d}t^2 = 0$$

或

$$1 - a^2 \left(\frac{\mathrm{d}t}{\mathrm{d}x}\right)^2 = 0. \tag{4.1.2}$$

由此定出两个方向:

$$\frac{\mathrm{d}t}{\mathrm{d}x} = \pm \frac{1}{a}, \tag{4.1.3}$$

称为**特征方向**. 解常微分方程 (4.1.3), 得到两族直线:

$$x - at = c_1, \quad x + at = c_2, \tag{4.1.4}$$

称为**特征**.

在研究波动方程的各种定解问题时, 特征起着重要作用. 例如, 我们用 u 沿特征的偏导数表示它沿 x, t 的偏导数, 则

$$\frac{\partial^2 u}{\partial t^2} = a^2 \left(\frac{\partial^2 u}{\partial c_1^2} - 2\frac{\partial^2 u}{\partial c_1 \partial c_2} + \frac{\partial^2 u}{\partial c_2^2}\right)$$
$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial c_1^2} + 2\frac{\partial^2 u}{\partial c_1 \partial c_2} + \frac{\partial^2 u}{\partial c_2^2}$$

于是方程 (4.1.1) 化为

$$\frac{\partial^2 u}{\partial c_1 \partial c_2} = 0,$$

从而其通解为

$$u = f_1(c_1) + f_2(c_2) = f_1(x - at) + f_2(x + at).$$

如果 u 在 x 轴的初值为

$$u(x,0) = \varphi_0(x), u_t(x,0) = \varphi_1(x), \quad -\infty < x < \infty, \tag{4.1.5}$$

则可以定出 f_1, f_2, 从而得

$$u(x,t) = \frac{1}{2}[\varphi_0(x+at) + \varphi_0(x-at)] + \frac{1}{2a}\int_{x-at}^{x+at} \varphi_1(\xi)\mathrm{d}\xi. \tag{4.1.6}$$

这是熟知的达朗贝尔 (d'Alembert) 公式.

公式 (4.1.6) 告诉我们, u 在点 (x_0,t_0) $(t_0 > 0)$ 的值仅依赖于初值函数 $\varphi_0(x)$, $\varphi_1(x)$ 在区间 $[x_0 - at_0, x_0 + at_0]$ 的值, 与区间外的初值无关, 故称 $[x_0 - at_0, x + at_0]$ 为点 (x_0,t_0) 的**依存域**. 其实, 区间 $[x_0 - at_0, x_0 + at_0]$ 上的初值不只确定 $u(x_0,t_0)$, 而且确定了 u 在以 $(x_0 - at_0, 0), (x_0 + at_0, 0), (x_0, t_0)$ 为顶点的三角形域内的值, 故称此三角形域为区间 $[x_0 - at_0, x_0 + at_0]$ 的**决定域**.

显然, 为了得到点 (x_0,t_0) 的依存域, 只需通过 (x_0,t_0) 作两条特征, 它们与 x 轴截出的闭区间即是. 为了得到区间 $[x_0 - at_0, x_0 + at_0]$ 的决定域, 过 $(x_0 - at_0, 0)$ 作第一特征 (斜率为正), 过 $(x_0 + at_0, 0)$ 作第二特征 (斜率为负), 它们交出的三角形域即为决定域 (参看图 1.1(a)).

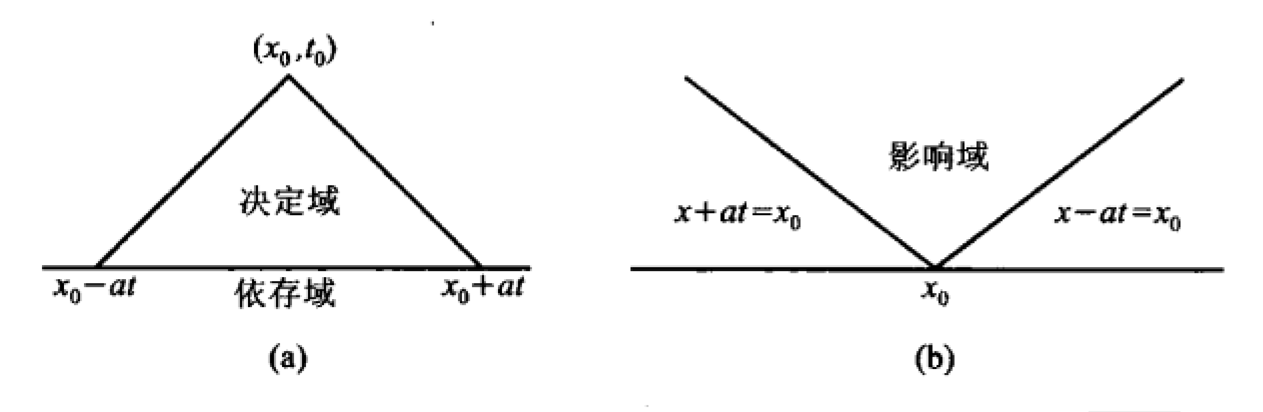


图 1.1

从公式 (4.1.6) 还看出, 对 x 轴上任一点 $(x_0,0)$, 依存域包含 $(x_0,0)$ 的一切 (x,t) 的集合恰好是以 $(x_0,0)$ 为顶点, 过 $(x_0,0)$ 的特征 $x - at = x_0$ 和 $x + at = x_0 (t > 0)$ 为边的角形域, 称之为 $(x_0,0)$ 的**影响域** (参看图 1.1(b)).

1.2 显格式

现在构造 (4.1.1) 的差分逼近. 取空间步长 h 和时间步长 τ, 用两族平行直线

$$x = x_j = jh, \quad j = 0, \pm1, \pm2, \cdots,$$

$$t = t_n = n\tau, \quad n = 0, 1, 2 \cdots$$

作矩形网格. 于网点 (x_j, t_n) 用 Taylor 展式, 得

$$\frac{u(x_{j+1},t_n)-2u(x_j,t_n)+u(x_{j-1},t_n)}{h^2}=u_{xx}(x_j,t_n)+\frac{h^2}{12}\frac{\partial^4}{\partial x^4}u(x_j,t_n)+O\left(h^4\right),$$
$$\frac{u(x_j,t_{n+1})-2u(x_j,t_n)+(x_j,t_{n-1})}{\tau^2}=u_{tt}(x_j,t_n)+\frac{\tau^2}{12}\frac{\partial^4}{\partial t^4}u(x_j,t_n)+O\left(\tau^4\right),$$

将 $u_{xx}(x_j,t_n)$, $u_{tt}(x_j,t_n)$ 代入波动方程, 得

$$\begin{aligned}&\frac{u(x_j,t_{n+1})-2u(x_j,t_n)+(x_j,t_{n-1})}{\tau^2}\\=&\frac{u(x_{j+1},t_n)-2u(x_j,t_n)+u(x_{j-1},t_n)}{h^2}+R_j^n(u),\end{aligned}$$

其中

$$R_j^n(u)=\frac{h^2}{12}\left(r^2\frac{\partial^4}{\partial t^4}u(x_j,t_n)-\frac{\partial^4}{\partial x^4}u(x_j,t_n)\right)+O\left(\tau^4+h^4\right) \tag{4.1.7}$$

是截断误差, $r=ah/\tau$ 是网比. 舍去 $R_j^n(u)$, 则得差分方程:

$$\frac{u_j^{n+1}-2u_j^n+u_j^{n-1}}{\tau^2}=a^2\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2},\quad j=0,\pm1,\pm2,\cdots,\quad n=1,2,\cdots, \tag{4.1.8}$$

这里 u_j^n 表示 u 于 (x_j,t_n) 的近似值. 初始条件用下列差分方程代替:

$$u_j^0=\varphi_0(x_j), \tag{4.1.9}$$

$$\frac{u_j^1-u_j^0}{\tau}=\varphi_1(x_j). \tag{4.1.10}$$

显然, (4.1.8) 的截断误差的阶是 $O(\tau^2+h^2)$, 而 (4.1.10) 的截断误差的阶仅为 $O(\tau)$. 为了提高精度, 也可以用中心差商代替 u_t, 得

$$\frac{u_j^1-u_j^{-1}}{2\tau}=\varphi_1(x_j). \tag{4.1.11}$$

于 (4.1.8) 令 $n=0$, 又得

$$\frac{u_j^1-2u_j^0+u_j^{-1}}{\tau^2}=a^2\frac{u_{j+1}^0-2u_j^0+u_{j-1}^0}{h^2}.$$

消去 u_j^{-1}, 则

$$u_j^1=\frac{r^2}{2}[\varphi_0(x_{j-1})+\varphi_0(x_{j+1})]+(1-r^2)\varphi_0(x_j)+\tau\varphi_1(x_j). \tag{4.1.12}$$

利用 (4.1.9), (4.1.10) (或 (4.1.12)) 可算出初始层 ($n=0$) 及第一层 ($n=1$) 各网格节点上的值. 然后利用 (4.1.8) 或

$$u_j^{n+1}=r^2(u_{j-1}^n+u_{j+1}^n)+2(1-r^2)u_j^n-u_j^{n-1}, \tag{4.1.13}$$

就可逐层算出任意网点的值.

(4.1.8) 是显的三层差分格式, 节点分布如图 1.2.

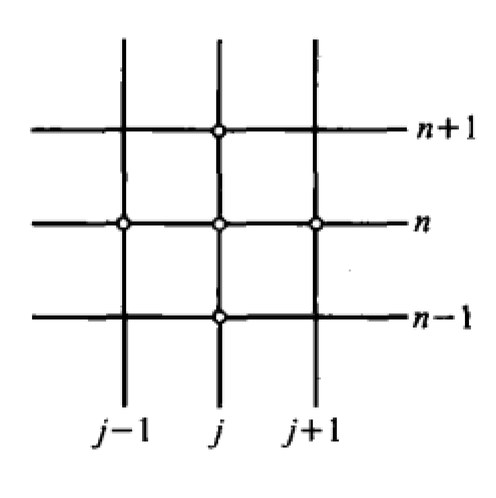


图 1.2

上述格式也可用以解混合问题:

$$\begin{cases}\dfrac{\partial^2 u}{\partial t^2}=a^2\dfrac{\partial^2 u}{\partial x^2}, & 0<x<l,\quad 0<t\leqslant T,\\ u(x,0)=\varphi_0(x), & u_t(x,0)=\varphi_1(x),\\ u(0,t)=\alpha(t), & u(l,t)=\beta(t).\end{cases} \tag{4.1.14}$$

这时取 $h=l/J$, $\tau=T/N$. 除 (4.1.8)—(4.1.10) 外, 再补充边值条件

$$u_0^n=\alpha(n\tau), u_J^n=\beta(n\tau). \tag{4.1.15}$$

1.3 稳定性分析

像抛物方程一样, 造出差分格式后, 要检验它是否稳定, 在什么条件下稳定, 这是差分方程理论的基本问题.

为了引用第三章判别稳定性的方法, 我们把波动方程 (4.1.1) 化成一阶微分方程组, 相应地把三层差分格式 (4.1.8) 化成两层差分格式. 一种简单的做法是引进变量 $v=\dfrac{\partial u}{\partial t}$, 于是 (4.1.1) 化为

$$\frac{\partial u}{\partial t}=v, \frac{\partial v}{\partial t}=a^2\frac{\partial^2 u}{\partial x^2}.$$

但对于构造差分逼近, 更常用的方法是引进变量 $w=a\dfrac{\partial u}{\partial x}$, 将 (4.1.1) 化为

$$\frac{\partial v}{\partial t}=a\frac{\partial w}{\partial x}, \frac{\partial w}{\partial t}=a\frac{\partial v}{\partial x}. \tag{4.1.16}$$

若令 $\boldsymbol{U}=(v,w)^{\mathrm{T}}$,

$$\boldsymbol{A}=\begin{pmatrix}0 & a\\ a & 0\end{pmatrix},$$

则 (4.1.16) 可写成为

$$\frac{\partial \boldsymbol{U}}{\partial t}-\boldsymbol{A}\frac{\partial \boldsymbol{U}}{\partial x}=0. \tag{4.1.17}$$

相应地, 将 (4.1.8) 写成等价的双层差分格式:

$$\begin{cases}\dfrac{v_j^{n+1}-v_j^n}{\tau}=a\dfrac{w_{j+\frac12}^n-w_{j-\frac12}^n}{h},\\[2ex] \dfrac{w_{j-\frac12}^{n+1}-w_{j-\frac12}^n}{\tau}=a\dfrac{v_j^{n+1}-v_{j-1}^{n+1}}{h},\end{cases} \tag{4.1.18}$$

其中

$$v_j^n=\frac{u_j^n-u_j^{n-1}}{\tau},\quad w_{j-\frac12}^{n+1}=a\frac{u_j^{n+1}-u_{j-1}^{n+1}}{h}.$$

现在用 Fourier 方法分析 (4.1.18) 的稳定性. 为此, 考虑具周期边值条件的混合问题. 按照第三章 §3, 以

$$v_j^n=V_1^n\exp(\mathrm{i}\alpha x_j),$$
$$w_j^n=V_2^n\exp(\mathrm{i}\alpha x_j)$$

$(\alpha=2p\pi/l,\mathrm{i}=\sqrt{-1})$ 代到 (4.1.17), 消去共同因子 $\exp(\mathrm{i}\alpha x_j)$ 和 $\exp\left(\mathrm{i}\alpha x_{j-\frac12}\right)$, 得

$$V_1^{n+1}-2\mathrm{i}r\left(\sin\frac{\pi ph}{l}\right)V_2^n=V_1^n,$$
$$-2\mathrm{i}r\left(\sin\frac{\pi ph}{l}\right)V_1^{n+1}+V_2^{n+1}=V_2^n,$$

或

$$\begin{pmatrix}V_1^{n+1}\\ V_2^{n+1}\end{pmatrix}=\boldsymbol{G}\left(\frac{\pi ph}{l}\right)\begin{pmatrix}V_1^n\\ V_2^n\end{pmatrix},$$

其中

$$\boldsymbol{G}\left(\frac{\pi ph}{l}\right)=\begin{pmatrix}1 & \mathrm{i}c\\ \mathrm{i}c & 1-c^2\end{pmatrix}\quad\left(c=2r\sin\frac{\pi ph}{l}\right) \tag{4.1.19}$$

为增长矩阵, $r=a\tau/h$ 为网比. 由第三章命题 4.1, 差分格式 (4.1.18) 稳定的充分必要条件是矩阵族

$$\{\boldsymbol{G}^n(\theta)\}\ (0\leqslant\theta\leqslant\pi,n=1,2\cdots)$$

一致有界.

注意 $\boldsymbol{G}(\theta)$ 的特征方程为

$$\lambda^2-(2-c^2)\lambda+1=0, \tag{4.1.20}$$

它的根按模 $\leqslant 1$ 的充要条件是 (见第三章 §4 习题 1)

$$|2-c^2| \leqslant 2,$$

即 $r\leqslant 1$. 这是差分格式稳定的必要条件 —— von Neumann 条件.

方程 (4.1.20) 的二根为

$$\lambda_{1,2}=\frac{1}{2}[(2-c^2)\pm \mathrm{i}|c|\sqrt{4-c^2}],$$

故

$$|\lambda_1-\lambda_2| = |c|\sqrt{4-c^2}.$$

又因 $c^2=4r^2\sin^2\theta\leqslant 4$, 故 $\lambda_2=\overline{\lambda}_1, |\lambda_1| = |\lambda_2| =1$, 从而

$$1-|\lambda_1| =0.$$

另一方面,

$$\boldsymbol{G}-\frac{1}{2}(\lambda_1+\lambda_2)\boldsymbol{I}=\begin{bmatrix}\frac{1}{2}c^2 & \mathrm{i}c\\ \mathrm{i}c & -\frac{1}{2}c^2\end{bmatrix},$$

其 F- 模为

$$\left\|\boldsymbol{G}-\frac{1}{2}(\lambda_1+\lambda_2)\boldsymbol{I}\right\|_{\mathrm{F}}=|c|\left(2+\frac{1}{2}c^2\right)^{\frac{1}{2}}.$$

为使第三章 §4 定理 4.3 推论中的条件 (β) 成立, 只需存在 $M>0$ 使

$$|c|\left(2+\frac{1}{2}c^2\right)^{\frac{1}{2}}\leqslant M|\lambda_1-\lambda_2| =M|c|\sqrt{4-c^2}. \tag{4.1.21}$$

上式当 $r<1$ 时显然成立. 若 $r=1$, 则当 $\theta=\dfrac{\pi}{2}$ 时, $c^2=4$, 上式右端为 0, 不等式 (4.1.21) 不成立, 故应要求 $r<1$.

总之, 差分格式 (4.1.18) 稳定的充要条件是网比

$$r=a\tau/h<1. \tag{4.1.22}$$

注 1.1 其实 $r=1$ 时, $\boldsymbol{G}(\theta)$ 于 $\theta=\dfrac{\pi}{2}$ 有重根 $\lambda_1=\lambda_2=-1$, 且初等因子的次数等于 2, 因此有相似变换 $\boldsymbol{S}$, 使

$$\boldsymbol{G}\left(\frac{\pi}{2}\right)=\boldsymbol{S}\begin{bmatrix}-1 & 1\\ 0 & -1\end{bmatrix}\boldsymbol{S}^{-1}$$

从而

$$G^n\left(\frac{\pi}{2}\right)=S\begin{bmatrix}(-1)^n & (-1)^{n-1}n\\ 0 & (-1)^n\end{bmatrix}S^{-1},\quad n=1,2,\cdots,$$

这表明 $G^n\left(\frac{\pi}{2}\right)$ 无界, 所以 $r=1$ 时格式 (4.1.18) 不稳定, 但它关于 n 是线性增长, 所以也称线性不稳定. 如果方程的解充分光滑, 则差分格式截断误差的阶达到 2, 按第二章定理 2.1 的证法仍可证明差分解收敛 (参看 [31]).

稳定性条件 (4.1.22) 有一直观几何解释. 从方程 (4.1.13) 看出, u_j^n 依赖前两层值 u_{j-1}^{n-1}, u_j^{n-1}, u_{j+1}^{n-1}, u_j^{n-2}, 这四个值又依赖 u_{j-2}^{n-2}, u_{j-1}^{n-2}, u_j^{n-2}, u_{j+1}^{n-2}, u_{j+2}^{n-2} 和 u_{j-1}^{n-3}, u_j^{n-3}, u_{j+1}^{n-3}. 依此类推, 可知 u_j^n 最终依赖初始层 $n=0$ 上的下列值:

$$u_{j-n}^0,\ u_{j-n+1}^0,\ \cdots,\ u_j^0,\cdots,\ u_{j+n-1}^0,\ u_{j+n}^0.$$

因此称 x 轴上含于区间 $[x_{j-n},x_{j+n}]$ 的网点为差分解 u_j^n 的依存域, 它是 x 轴上被过 (x_j,t_n) 的两条直线

$$x-x_j=\pm\frac{h}{\tau}(t-t_n)$$

截下的区间所覆盖的网域. 注意过 (x_j,t_n) 的两条特征为 $x-x_j=\pm a(t-t_n)$, 差分方程稳定性的必要条件为 $r\leqslant 1$ 或

$$\tau/h\leqslant a^{-1}.$$

可见差分方程稳定的必要条件是: 差分解的依存域必须包含微分方程解的依存域, 否则, 差分方程不稳定.

现在利用依存域的概念证明: 当 $r>1$ 时差分解不收敛. 如图 1.3, $r>1$, 微分方程解的依存域 $[P',Q']$ 大于差分解的依存域 $[P,Q]$. 固定 (x_j,t_n), 让网格步长变小, 但网比 r 保持不变, 则依存域 $[P',Q'],[P,Q]$ 不变. 显然, 若改变区间 (P',P) 和 (Q,Q') 上的初值, 但 $[P,Q]$ 上的初值不变, 则 $u(x_j,t_n)$ 可取不同值, 而 u_j^n 当 $h\to 0,\tau\to 0$ 时 (r 固定不变) 是一串确定的数列, 它不可能收敛到不同的 $u(x_j,t_n)$.

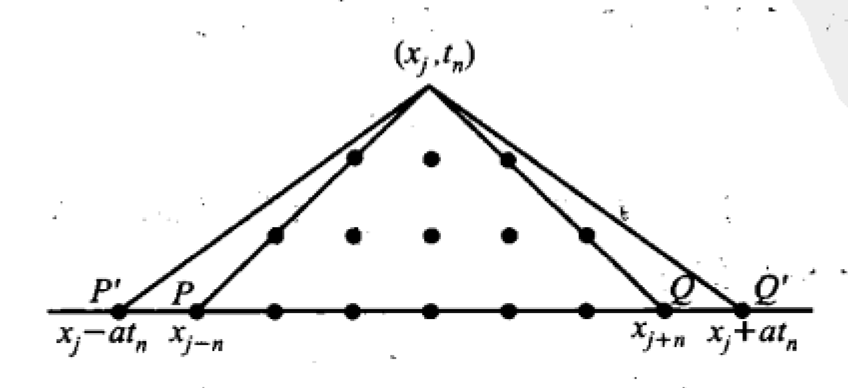


图 1.3

总之我们知道, 当 $r<1$ 时, 差分方程稳定, 因而差分解收敛 (参看第三章 §2). Courant 等曾证明 $r=1$ 时差分解仍收敛, 但要求有更光滑的初值 (参看注 1.1 及 [31]). 习惯上也称 $r\leqslant 1$ 为 Courant 条件或 CFL 条件 (Courant-Friedrichs-Lewy condition).

1.4 隐格式

为了得到恒稳定的差分格式, 用第 $n-1$ 层、n 层、$n+1$ 层的中心差商的权平均去逼近 u_{xx} 得到下列差分格式:

$$\frac{u_j^{n+1}-2u_j^n+u_j^{n-1}}{\tau^2}=a^2\left[\theta\frac{u_{j+1}^{n+1}-2u_j^{n+1}+u_{j-1}^{n+1}}{h^2}+\right.$$
$$\left.(1-2\theta)\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2}+\theta\frac{u_{j+1}^{n-1}-2u_j^{n-1}+u_{j-1}^{n-1}}{h^2}\right], \tag{4.1.23}$$

其中 $0\leqslant\theta\leqslant 1$ 是参数. 当 $\theta=0$ 时就是显格式 (4.1.8). 实际有兴趣的参数是 $\theta=\frac{1}{4}$, 此时差分格式可化为

$$\begin{cases}\dfrac{v_j^{n+1}-v_j^n}{\tau}=a\dfrac{w_{j+\frac12}^n-w_{j-\frac12}^n+w_{j+\frac12}^{n+1}-w_{j-\frac12}^{n+1}}{2h},\\[2ex] \dfrac{w_{j-\frac12}^{n+1}-w_{j-\frac12}^n}{\tau}=a\dfrac{v_j^{n+1}-v_{j-1}^{n+1}+v_j^n-v_{j-1}^n}{2h}.\end{cases} \tag{4.1.24}$$

其增长矩阵为

$$\boldsymbol{G}(\theta)=\begin{pmatrix}\dfrac{1-c^2/4}{1+c^2/4} & \dfrac{\mathrm{i}c}{1+c^2/4}\\[2ex] \dfrac{\mathrm{i}c}{1+c^2/4} & \dfrac{1-c^2/4}{1+c^2/4}\end{pmatrix},\quad c=2r\sin\theta.$$

可以证明, $\boldsymbol{G}(\theta)$ 的特征值按绝对值等于 1, 且 $\boldsymbol{G}$ 是酉矩阵. 因此 $\boldsymbol{G}$ 的欧氏模 $\|\boldsymbol{G}\|=1$, 从而矩阵族 $\{\boldsymbol{G}^n(\theta)\}$ 一致有界, 故 (4.1.22) 恒稳定.

1.5 数值例子

求解

$$\begin{aligned}&u_{tt}=u_{xx},\quad 0<x<1,\quad t>0,\\ &u(0,t)=u(1,t)=0,\quad t>0,\\ &u(x,0)=\sin 4\pi x, u_t(x,0)=\sin 8\pi x,\quad 0<x<1.\end{aligned}$$

(精确解: $u=\cos 4\pi t\sin 4\pi x+(\sin 8\pi t\sin 8\pi x)/8\pi$)
取空间步长 $h=1/J$, 时间步长 $\tau>0$, 网比 $r=\tau/h$.

显格式为

$$\begin{aligned}&\frac{u_j^{n+1}-2u_j^n+u_j^{n-1}}{\tau^2}=\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2},\\&u_0^n=u_J^n=0,\\&u_j^0=\sin 4\pi x_j, u_j^1=\sin 4\pi x_j+\tau\sin 8\pi x_j\end{aligned}$$

方案 I　$h=1/400=0.002\ 5$, $\tau=1/500=0.002$, 此时 $r=4/5$. 计算 $t=1,2,3,4,5$ 的差分解.

方案 II　$h=\tau=1/400=0.002\ 5$, 此时 $r=1$. 计算 $t=1,2,3,4,5$ 的差分解.

表 1.1 列出方案 I, II 的差分解的误差阶. 从表中看出, 方案 II 的精度比 I 高很多, 这是因为格式当 $r=1$ 时截断误差有最高阶 $O(h^4)$, $r\neq 1$ 时截断误差的阶为 $O(h^2)$ (参看节 1.2 的 (4.1.7)), 且时间层数受条件 $n\tau\leqslant 5$ 限制.

表 1.1

方案	x \ t	1	2	3	4	5
I	误差阶	1.0×10^{-4}	1.0×10^{-3}	1.0×10^{-3}	1.0×10^{-3}	1.0×10^{-3}
II		1.0×10^{-13}	1.0×10^{-13}	1.0×10^{-13}	1.0×10^{-13}	1.0×10^{-13}

习　　题

1. 就二维波动方程导出显格式, 并给出稳定性条件.

2. 证明格式 (4.1.24) 恒稳定.

3. 取初值 $v_j^0=(-1)^j, w_{j+\frac{1}{2}}=0$, 网比 $r=a\tau/h=1$. 求差分方程 (4.1.18) 的解, 并用计算机验证. (答案: $v_j^n=(-1)^{n+1}(1-2n), w_{j+\frac{1}{2}}^n=(-1)^{n+1}2n$.)

4. (实习题) 利用差分格式 (4.1.8)–(4.1.10) 求下列波动方程混合边值问题的解:

$$\begin{cases}\dfrac{\partial^2 u}{\partial t^2}-\dfrac{\partial^2 u}{\partial x^2}=0, & 0<x<1, t>0,\\ u|_{t=0}=\sin\pi x,\ \left.\dfrac{\partial u}{\partial t}\right|_{t=0}=\cos\pi x, & 0<x<1,\\ u(0,t)=u(1,t)=0, & t\geqslant 0.\end{cases}$$

(精确解 $u=\sin\pi(x-t)+\sin\pi(x+t)$)

(1) 取 $\tau=0.05$, $h=0.1$, 计算 $t=0.5$, 1.0, 1.5, 2.0 的解.

(2) 取 $\tau=h=0.1$, 计算 $t=0.5, 1.0, 1.5, 2.0$ 的解.

5. (实习题) 利用差分格式 (4.1.8)–(4.1.9), (4.1.12) 重复题 4 的计算, 比较计算结果.

§2 一阶线性双曲方程组

从本节起, 我们讨论一阶线性双曲方程组的差分解法. 由于构造差分格式与偏微分方程的特征及解的性质有关, 所以在讨论数值解法之前, 先回顾一下双曲方程组的某些基本概念 (参看 [3]).

2.1 双曲型方程组及其特征

设有含 n 个未知函数 $\boldsymbol{u} = (u_1(x,t), \cdots, u_n(x,t))$ 和 n 个方程的一阶线性偏微分方程组:

$$L_i(\boldsymbol{u}) = \sum_{j=1}^{n} b_{ij} \frac{\partial u_j}{\partial t} + \sum_{j=1}^{n} a_{ij} \frac{\partial u_j}{\partial x} = c_i, \quad i = 1, 2, \cdots, n, \tag{4.2.1}$$

其中 $b_{ij} = b_{ij}(x,t), a_{ij} = a_{ij}(x,t)$ 及 $c_i = c_i(x,t)$ 都是域 G 上的光滑函数. 采用矩阵、向量记号

$$\boldsymbol{B} = (b_{ij})_{n\times n}, \boldsymbol{A} = (a_{ij})_{n\times n},$$
$$\boldsymbol{c} = (c_1, c_2, \cdots, c_n)^{\mathrm{T}},$$

将 (4.2.1) 写成

$$L(\boldsymbol{u}) = \boldsymbol{B}\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{A}\frac{\partial \boldsymbol{u}}{\partial x} = \boldsymbol{c}. \tag{4.2.2}$$

假定矩阵 $\boldsymbol{B}$ 有逆, 则不失一般性地可设 $\boldsymbol{B} = \boldsymbol{I}$ (单位矩阵), 只考虑如下形式的方程组:

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{A}\frac{\partial \boldsymbol{u}}{\partial x} = \boldsymbol{c}. \tag{4.2.3}$$

定义 2.1 我们说 (4.2.3) 于点 $(x,t) \in G$ 是 **(狭义)** 双曲方程组, 如果矩阵 $\boldsymbol{A} = \boldsymbol{A}(x,t)$ 有 n 个实的互异特征值

$$\lambda_1(x,t) < \lambda_2(x,t) < \cdots < \lambda_n(x,t). \tag{4.2.4}$$

假若 (4.2.3) 于每一点 $(x,t) \in G$ 为双曲, 则说它是 G 上的双曲方程组.

本节总设 (4.2.3) 是双曲方程组, 行向量

$$\boldsymbol{l}^{(1)}(x,t), \cdots, \boldsymbol{l}^{(n)}(x,t) \tag{4.2.5}$$

是矩阵 $\boldsymbol{A}$ 相应于 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 的左特征向量系, 即

$$\boldsymbol{l}^{(i)}\boldsymbol{A} = \lambda_i \boldsymbol{l}^{(i)}, \quad i = 1, 2, \cdots, n.$$

用 $\boldsymbol{l}^{(i)}$ 左乘 (4.2.3) 两端, 得

$$\boldsymbol{l}^{(i)} u_t + \lambda_i \boldsymbol{l}^{(i)} u_x = \boldsymbol{l}^{(i)} \boldsymbol{c}, \quad i = 1, 2, \cdots, n. \tag{4.2.6}$$

设 $\boldsymbol{l}^{(i)} = (l_1^{(i)}, l_2^{(i)}, \cdots, l_n^{(i)})$, 则 (4.2.6) 即

$$\sum_{j=1}^{n} l_j^{(i)} \left(\frac{\partial u_j}{\partial t} + \lambda_i \frac{\partial u_j}{\partial x} \right) = \sum_{j=1}^{n} l_j^{(i)} c_j, \quad i = 1, 2, \cdots, n. \tag{4.2.7}$$

在 x, t 平面域 G 内各点作 n 个方向

$$\mathrm{d}t : \mathrm{d}x = 1 : \lambda_i \tag{4.2.8}$$

或

$$\tau_i : \frac{\mathrm{d}x}{\mathrm{d}t} = \lambda_i, \quad i = 1, 2, \cdots, n, \tag{4.2.9}$$

则沿 τ_i,

$$\left(\frac{\mathrm{d}u_j}{\mathrm{d}t} \right)_{\tau_i} = \frac{\partial u_j}{\partial t} + \lambda_i \frac{\partial u_j}{\partial x}.$$

于是将 (4.2.7) 化成常微分方程组:

$$\sum_{j=1}^{n} l_j^{(i)} \left(\frac{\mathrm{d}u_j}{\mathrm{d}t} \right)_{\tau_i} = \sum_{j=1}^{n} l_j^{(i)} c_j, \quad i = 1, 2, \cdots, n. \tag{4.2.10}$$

由 (4.2.9) 确定的 n 个不同方向称为**特征方向**; 由特征方向确定的 n 族曲线, 称为**特征曲线**, 简称**特征**. 沿每一特征, 方程组 (4.2.3) 化成常微分方程组 (4.2.10), 称为原方程的**特征关系**. 在特征上成立的特征关系乃是利用特征概念研究双曲方程的基础.

特征关系 (4.2.10) 的每一方程, 只出现未知函数沿同一方向的导数. 实际上还可进一步化简, 使每个方程只出现一个函数的方向导数. 因

$$l_j^{(i)} \left(\frac{\mathrm{d}u_j}{\mathrm{d}t} \right)_{\tau_i} = \left(\frac{\mathrm{d}l_j^{(i)} u_j}{\mathrm{d}t} \right)_{\tau_i} - u_j \left(\frac{\mathrm{d}l_j^{(i)}}{\mathrm{d}t} \right)_{\tau_i},$$

若令

$$r_i = r_i(x, t) = \sum_{j=1}^{n} l_j^{(i)} u_j, \tag{4.2.11}$$

则可将 (4.2.10) 化成

$$\left(\frac{\mathrm{d}r_i}{\mathrm{d}t}\right)_{\tau_i} = \frac{\partial r_i}{\partial t} + \lambda_i \frac{\partial r_i}{\partial x} = \sum_{j=1}^{n}\left(l_j^{(i)} c_i + u_j \left(\frac{\mathrm{d}l_j^{(i)}}{\mathrm{d}t}\right)_{\tau_i}\right).$$

由于 (4.2.11) 右端系数矩阵 $\boldsymbol{L} = (l_j^{(i)})_{n\times n}$ 有逆 $\boldsymbol{L}^{-1} = \boldsymbol{M} = (m_{ij})_{n\times n}$, 故 $\boldsymbol{u}$ 可通过 $\boldsymbol{r} = (r_1, r_2, \cdots, r_n)^{\mathrm{T}}$ 表为

$$u_j = \sum_{k=1}^{n} m_{jk} r_k, \quad j = 1, 2\cdots, n.$$

于是

$$\left(\frac{\mathrm{d}r_i}{\mathrm{d}t}\right)_{\tau_i} = g_i = \sum_{j=1}^{n} l_j^{(i)} c_j + \sum_{j=1}^{n}\left(\frac{\mathrm{d}l_j^{(i)}}{\mathrm{d}t}\right)_{\tau_i} \sum_{k=1}^{n} m_{jk} r_k, \quad i = 1, 2, \cdots, n. \tag{4.2.12}$$

这样就把 (4.2.10) 化成了对角方程组 (4.2.12). 新变量 $r_i (i = 1, 2, \cdots, n)$ 称为 **Riemann 不变量**.

当系数矩阵 $\boldsymbol{A}$ 为常矩阵 (因而 $l_j^{(i)}$ 与 x, t 无关), 且右端 $c_i = 0$ 时,(4.2.12) 简化为

$$\left(\frac{\mathrm{d}r_i}{\mathrm{d}t}\right)_{\tau_i} = 0.$$

可见 r_i 沿特征

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \lambda_i, \quad i = 1, 2, \cdots, n$$

是常量.

例 2.1 波动方程 (4.1.1) 可化成一阶方程组:

$$\frac{\partial \boldsymbol{u}}{\partial t} - \boldsymbol{A}\frac{\partial \boldsymbol{u}}{\partial x} = \boldsymbol{0}, \tag{4.2.13}$$

$$\boldsymbol{A} = \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix},$$

其中 $\boldsymbol{u} = (v, w)^{\mathrm{T}}$. $-\boldsymbol{A}$ 的特征方程为 $\lambda^2 - a^2 = 0$, 有二互异实根 $\pm a$, 故 (4.2.13) 是双曲方程. 特征是两族直线 $x - at = c_1, x + at = c_2$. 特征关系为

$$\frac{\mathrm{d}v}{\mathrm{d}t} - \frac{\mathrm{d}w}{\mathrm{d}t} = 0, \quad 沿\ x - at = c_1$$
$$\frac{\mathrm{d}v}{\mathrm{d}t} + \frac{\mathrm{d}w}{\mathrm{d}t} = 0, \quad 沿\ x + at = c_2.$$

由 (4.2.11) 定义的 Riemann 不变量为

$$r_1 = v - w,\ r_2 = v + w.$$

于是 (4.2.13) 可写成 Riemann 不变量形式:

$$\frac{\partial r_1}{\partial t} + a\frac{\partial r_1}{\partial x} = 0, \quad \frac{\partial r_2}{\partial t} - a\frac{\partial r_2}{\partial x} = 0.$$

显然 r_1 沿 $x - at = c_1$ 等于常数, r_2 沿 $x + at = c_2$ 等于常数, 因此

$$r_1 = f(x - at), \quad r_2 = g(x + at).$$

从而

$$\begin{aligned} v &= \frac{1}{2}(f(x-at) + g(x+at)), \\ w &= -\frac{1}{2}(f(x-at) - g(x+at)). \end{aligned}$$

例 2.2 考虑在静止气体中小扰动 (声音) 传播所满足的方程组:

$$\begin{cases} \dfrac{\partial u}{\partial t} + \dfrac{c_0^2}{\rho_0}\dfrac{\partial \rho}{\partial x} = 0, \\ \dfrac{\partial \rho}{\partial t} + \rho_0\dfrac{\partial u}{\partial x} = 0 \end{cases} \tag{4.2.14}$$

其中 u 和 ρ 分别表示扰动后的质点速度及密度, ρ_0 及 c_0 表示静止气体的密度和音速 (ρ_0, c_0 都是正常数).

现在

$$\boldsymbol{A} = \begin{pmatrix} 0 & \dfrac{c_0^2}{\rho_0} \\ \rho_0 & 0 \end{pmatrix},$$

其特征值为 $\pm c_0$, 所以 (4.2.14) 是双曲方程. 特征是两族直线 $x - c_0 t = c_1, x + c_0 t = c_2$. 特征关系为

$$\begin{aligned} \rho_0\frac{\mathrm{d}u}{\mathrm{d}t} + c_0\frac{\mathrm{d}\rho}{\mathrm{d}t} &= 0, \quad 沿\ x - c_0 t = c_1, \\ \rho_0\frac{\mathrm{d}u}{\mathrm{d}t} - c_0\frac{\mathrm{d}\rho}{\mathrm{d}t} &= 0, \quad 沿\ x + c_0 t = c_2. \end{aligned}$$

Riemann 不变量为

$$r_1 = \rho_0 u + c_0 \rho, \quad r_2 = \rho_0 u - c_0 \rho.$$

r_1 沿 $x - c_0 t = c_1$ 等于常数, r_2 沿 $x + c_0 t = c_2$ 等于常数.

2.2 Cauchy 问题、依存域、影响域和决定域

考虑双曲方程组 (4.2.3) 的如下 Cauchy 问题: 在线段

$$\overline{P_1 P_2}: t = 0, \quad a \leqslant x \leqslant b$$

的一邻域内, 求 (4.2.3) 的解 $\boldsymbol{u}(x,t)=(u_1(x,t),\cdots,u_n(x,t))^{\mathrm{T}}$, 在 $\overline{P_1P_2}$ 上取给定的初值 $\boldsymbol{u}^0(x)$, 即

$$\boldsymbol{u}(x,0)=\boldsymbol{u}^0(x), x\in\overline{P_1P_2}\,. \tag{4.2.15}$$

由于 x 轴的方向 $(\mathrm{d}x,\mathrm{d}t)=(1,0)$ 不是特征方向 (因 $-\infty<\lambda_i<\infty, i=1,2,\cdots,n$), 所以 Cauchy 问题适定, 即在 $\overline{P_1P_2}$ 一邻域内有唯一解, 且解连续依赖初值 (参看 [3]).

现在讨论解与初值的关系. 我们称由

$$\frac{\mathrm{d}x}{\mathrm{d}t}=\lambda_i(x,t),\quad i=1,2,\cdots,n$$

确定的曲线族为第 i 族特征. 由常微分方程理论知道, 同族特征不相交, 所以过任一点恰有 n 个不同特征经过. 又 x 轴 $(t=0)$ 方向不是特征方向, 对任一点 $P=(x,t)(t>0)$, 取过 P 的两条特征

$$\begin{aligned}\tau_n:\frac{\mathrm{d}x}{\mathrm{d}t}&=\lambda_n(x,t)=\lambda_{\max},\\ \tau_1:\frac{\mathrm{d}x}{\mathrm{d}t}&=\lambda_1(x,t)=\lambda_{\min},\end{aligned}$$

假定它们与 x 轴依次交于 P_1 和 P_2. 曲线段 $\overline{PP_1}$ 和 $\overline{PP_2}$ 与直线段 $\overline{P_1P_2}$ 围成一曲边三角形, 如图 2.1(a) 所示. 显然过 P 的 n 条特征随 t 递减, 都和 x 轴交于 $\overline{P_1P_2}$. 现在从 P 出发沿特征 (按 t 减小方向) 积分常微分方程组 (4.2.12), 则知 Riemann 不变量 $\boldsymbol{r}=(r_1,r_2,\cdots,r_n)$, 从而 $u=u(x,t)$ 于 P 的值只和 $\overline{P_1P_2}$ 上的初值有关. 称这样的线段为点 P 的**依存域**. 又曲边三角形 P_1PP_2 围成的区域 G 内任一点的 $u(x,t)$ 也由 $\overline{P_1P_2}$ 上的初值唯一决定, 故 G 称为 $\overline{P_1P_2}$ 的**决定域**. 同时也看出, 线段 $\overline{Q_1Q_2}$ (图 2.1(b)) 上的初值, 随 t 增加可影响到上半平面过 Q_1 的第一特征与过 Q_2 的第 n 特征之间的区域中任一点, 但不影响其他点, 故称之为 $\overline{Q_1Q_2}$ 的**影响域**.

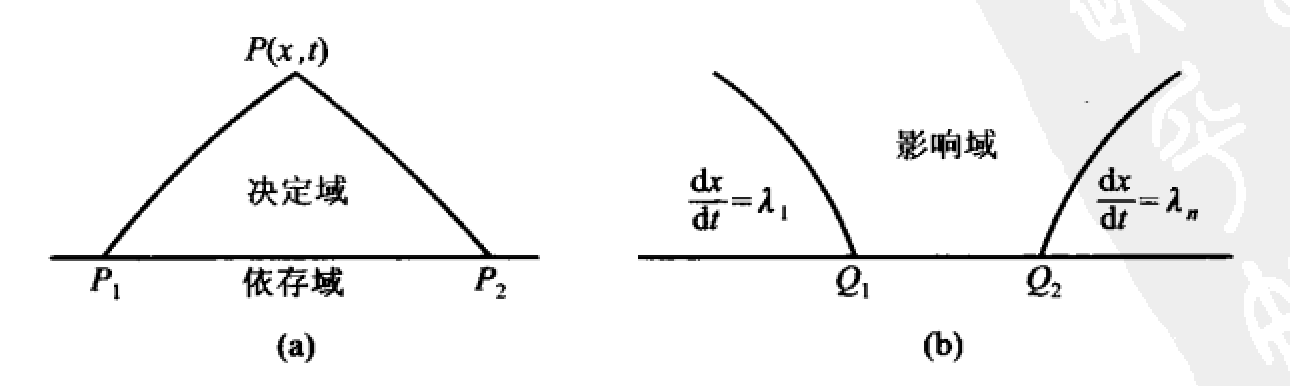


图 2.1

2.3　初边值问题

为简单计, 我们只讨论含两个未知函数 u, v 的如下对角形方程组:

$$\left\{\begin{array}{l}\dfrac{\partial u}{\partial t}+\lambda_1\dfrac{\partial u}{\partial x}=f_1,\\[2mm] \dfrac{\partial v}{\partial t}+\lambda_2\dfrac{\partial v}{\partial x}=f_2,\end{array}\right. \tag{4.2.16}$$

其中 $\lambda_1<\lambda_2, f_i=a_iu+b_iv+c_i, i=1,2$. 此时平面上有两族特征, 我们规定特征的正向是指向 t 增加的方向. 过 (x,t) 的任一非特征方向 α 说是时向的, 如果 α 或 $-\alpha$ 介于过此点的两正特征方向之间, 否则就说 α 是空向的 (参看图 2.2).

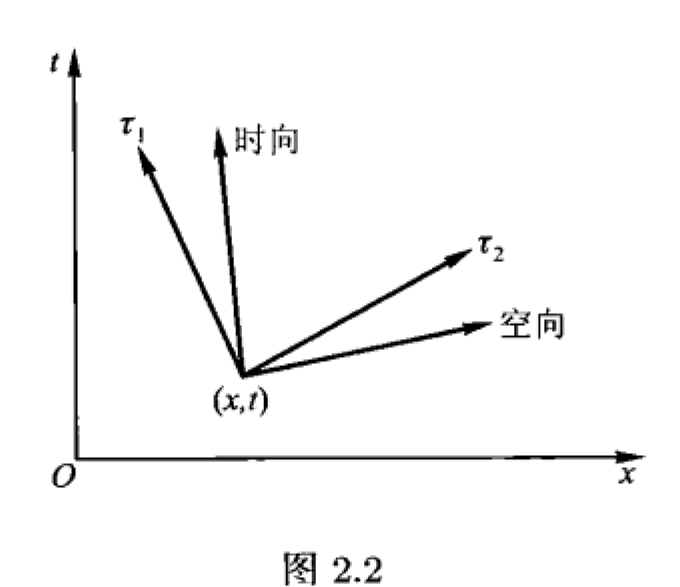


图 2.2

在 $t>0$ 和 $0<x<l$ 求方程 (4.2.16) 的解 u, v, 满足初始条件

$$u(x,0)=u^0(x),\quad v(x,0)=v^0(x),\quad 0\leqslant x\leqslant l \tag{4.2.17}$$

和适当的边值条件, 使解存在、唯一. 如何给边值条件才是恰当的? 这里我们只作一粗略但颇有启发性的分析, 严格论证就不给了, 有兴趣的读者可参看偏微分方程的专门著作 (例如 [3]).

先设方程的系数是常数, 且右端 $f_i=0\ (i=1,2)$. 此时两族特征都是直特征, 方程 (4.2.16) 可积分, 得

$$\begin{aligned}u(x,t)=&\ \text{常数, 当 } x-\lambda_1t=c_1,\\ v(x,t)=&\ \text{常数, 当 } x-\lambda_2t=c_2.\end{aligned}$$

设 $G: 0<x<l, t>0$. 分三种情形 (参看图 2.3):

(i) $\lambda_2>\lambda_1>0$. 此时随时间 t 递增, 两族特征由左边界 $(x=0)$ 进入 G, 由右边界 $(x=l)$ 离开 G (垂线 $x=0$ 和 $x=l$ 是空向). u, v 在 G 的性质与左边值有关, 与右边值无关, 故 u, v 的边值条件应给在左端点, 即

$$u(0,t)=u_0(t),\quad v(0,t)=v_0(t).$$

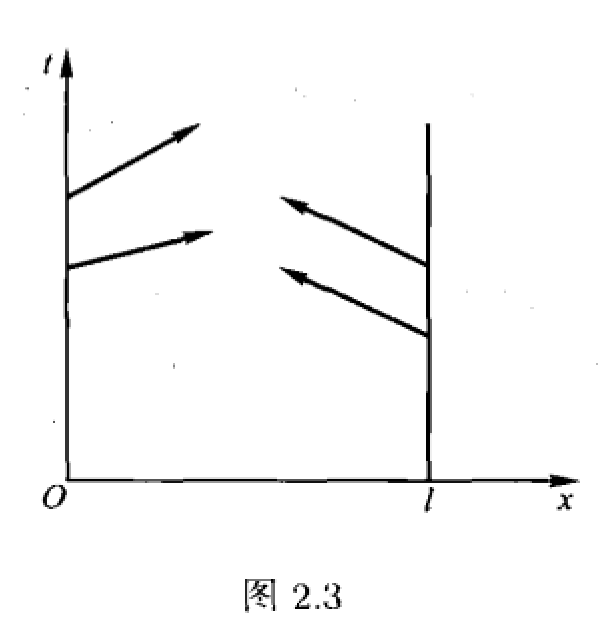


图 2.3

(ii) $\lambda_1 < \lambda_2 < 0$. 与 (i) 类似 (垂线 $x=0$ 和 $x=l$ 也是空向), u 和 v 的边值条件应给在右端点, 即

$$u(l,t)=u_1(t), \quad v(l,t)=v_1(t).$$

(iii) $\lambda_1 < 0 < \lambda_2$, 此时第一族特征由右边界进入 G, 第二族特征由左边界进入 G (垂线 $x=0$ 和 $x=l$ 是时向), 故边值条件应给成:

$$u(l,t)=u_1(t), \qquad v(0,t)=v_0(t).$$

虽然上述结论是就常系数和齐右端情形给出的, 但对变系数和非齐右端也成立. 此外, 若双曲方程不是对角形的, 则可用节 2.1 的方法化为对角形方程组 (4.2.12), 然后就三种不同情形对 Riemann 不变量 $r_i(i=1,2)$ 给边值条件.

习　题

1. 试求下列初边值问题的解:

$$\begin{cases} \dfrac{\partial u}{\partial t}+\dfrac{\partial u}{\partial x}=0, & 0<x<\infty, t>0, \\ u(x,0)=|x-1|, & u(0,t)=1. \end{cases}$$

2. 试求下列初边值问题的解:

$$\begin{cases} \dfrac{\partial u}{\partial t}+(1+x)\dfrac{\partial u}{\partial x}=0, & 0<x<\infty, \quad t>0, \\ u(x,0)=\varphi(x), & u(0,t)=1, \quad t\geqslant 0, \end{cases}$$

其中

$$\varphi(x)=\begin{cases} 1, & 0\leqslant x<1, \\ 0, & x\geqslant 1. \end{cases}$$

§3　初值问题的差分逼近

双曲方程与椭圆方程、抛物方程的一个重要区别, 是双曲方程具有特征和特征关系, 其解对初值有局部依赖性质. 初值的函数性质 (如间断、弱间断等) 将沿特征传播, 因而解一般没有光滑性. 在构造双曲方程的差分逼近时, 应考虑这些特性. 迄今已发展了许多逼近双曲方程的差分格式, 这里只介绍常见的几种, 有兴趣的读者可参看文献 [9]、[15]、[28]、[31].

3.1　迎风格式

首先考虑线性常系数方程式:

$$\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=0. \tag{4.3.1}$$

这个方程虽简单, 但对我们构造差分格式很有启发. 我们的主要目的是构造差分格式, 因此先讨论纯初值问题, 然后在 §4 对初边值问题作若干注记.

沿用 §1 的记号, 作 (4.3.1) 的差分逼近. 按照差商代替微商的办法, 自然有如下三种格式:

$$\frac{u_j^{n+1}-u_j^n}{\tau}+a\frac{u_j^n-u_{j-1}^n}{h}=0, \tag*{$(4.3.2)_1$}$$

$$\frac{u_j^{n+1}-u_j^n}{\tau}+a\frac{u_{j+1}^n-u_j^n}{h}=0, \tag*{$(4.3.2)_2$}$$

$$\frac{u_j^{n+1}-u_j^n}{\tau}+a\frac{u_{j+1}^n-u_{j-1}^n}{2h}=0. \tag*{$(4.3.2)_3$}$$

前两个方程截断误差的阶为 $O(\tau+h)$, 第三个方程是 $O(\tau+h^2)$.

从稳定性分析将会知道, 这三个格式并不都可用. 记

$$r=a\tau/h. \tag{4.3.3}$$

将 $(4.3.2)_1\sim(4.3.2)_3$ 改写成

$$u_j^{n+1}=ru_{j-1}^n+(1-r)u_j^n, \tag*{$(4.3.2)_1'$}$$

$$u_j^{n+1}=(1+r)u_j^n-ru_{j+1}^n, \tag*{$(4.3.2)_2'$}$$

$$u_j^{n+1}=u_j^n+\frac{r}{2}u_{j-1}^n-\frac{r}{2}u_{j+1}^n. \tag*{$(4.3.2)_3'$}$$

按 Fourier 方法, 以 $u_j^n = v^n \exp(\mathrm{i}\alpha x_j)$ ($\mathrm{i} = \sqrt{-1}, x_j = jh$, α 是任意实参数) 代到上述方程, 消去公因子, 分别得

$$v^{n+1} = (r\mathrm{e}^{-\mathrm{i}\alpha h} + (1-r))v^n = \lambda_1 v^n,$$
$$v^{n+1} = ((1+r) - r\mathrm{e}^{\mathrm{i}\alpha h})v^n = \lambda_2 v^n,$$
$$v^{n+1} = (1 - \mathrm{i}r\sin\alpha h)v^n = \lambda_3 v^n.$$

因为 $|\lambda_3| = \sqrt{1 + r^2\sin^2\alpha h}$ 对任何 $r \neq 0$ 都不满足 von Neumann 条件, 故 $(4.3.2)_3$ 恒不稳定. 其次, $|\lambda_1| \leqslant 1$ 等价于 $r^2 \leqslant r$, 即

$$(a\tau/h)^2 \leqslant (a\tau/h),$$

故 $(4.3.2)_1$ 稳定的充要条件是

$$a \geqslant 0, \left|\frac{a\tau}{h}\right| \leqslant 1. \tag{4.3.4}$$

同理, $(4.3.2)_2$ 稳定的充要条件是

$$a \leqslant 0, \left|\frac{a\tau}{h}\right| \leqslant 1. \tag{4.3.5}$$

这说明 $a \geqslant 0$ 时只有 $(4.3.2)_1$ 可用, $a \leqslant 0$ 时只有 $(4.3.2)_2$ 可用.

现在我们用特征性质说明这些格式及其稳定性条件. 注意 (4.3.1) 的特征为

$$\mathrm{d}x/\mathrm{d}t = a,$$

特征关系是

$$\mathrm{d}u/\mathrm{d}t = 0.$$

设已知 $u_{j+1}^n, u_j^n, u_{j-1}^n$, 要造出 u_j^{n+1} 的计算公式. 如图 3.1, 过 $P_0(j, n+1) = (jh, (n+1)\tau)$ 作特征, 斜率为 $\mathrm{d}t/\mathrm{d}x = 1/a$. 当 $a > 0$ 时, 特征偏左, 与直线 $t = t_n = n\tau$ 的交点 Q 位于 $Q_0(j,n)$ 左侧. u 沿线段 $\overline{P_0Q}$ 等于常数, 故 $u_j^{n+1} = u_{P_0} = u_Q$. 利用 Q_{-1}, Q_0 作线性插值, 得

$$\begin{aligned} u_j^{n+1} = u_Q &\approx (u_{Q_{-1}} \cdot \overline{QQ_0} + u_{Q_0}(h - \overline{QQ_0}))/h \\ &= (u_{j-1}^n \cdot a\tau + u_j^n(h - a\tau))/h \\ &= ru_{j-1}^n + (1-r)u_j^n. \end{aligned}$$

这就是 $(4.3.2)_1'$ 或 $(4.3.2)_1$. 稳定性条件 (4.3.4) 的第二个不等式意味着 Q 应落在 Q_{-1}, Q_0 之间, 即差分方程的依存域包含微分方程的依存域. 类似的解释也适用于 $a < 0$, 此时过 P_0 的特征偏右, 与 $t = t_n$ 的交点 Q 落在 Q_0 右侧. 利用

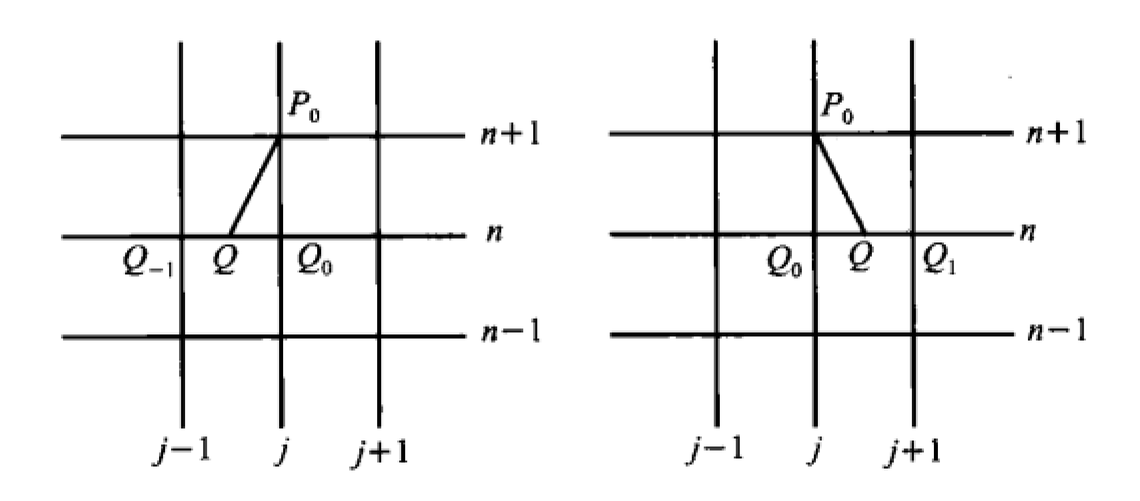


图 3.1

Q_0, Q_1, 作线性插值便得到 u_Q, 从而得出 u_j^{n+1}, 这就是 $(4.3.2)_2'$ 或 $(4.3.2)_2$. 这说明差分格式 $(4.3.2)_1$, $(4.3.2)_2$ 与特征走向有内在联系.

用同样思想可构造变系数方程式

$$\frac{\partial u}{\partial t} + a(x)\frac{\partial u}{\partial x} = 0$$

的差分格式. 此时 a 可能变号, 因此相应的格式为

$$(4.3.6) \qquad \begin{cases} \dfrac{u_j^{n+1} - u_j^n}{\tau} + a_j \dfrac{u_j^n - u_{j-1}^n}{h} = 0, & \text{若 } a_j \geqslant 0, \\ \dfrac{u_j^{n+1} - u_j^n}{\tau} + a_j \dfrac{u_{j+1}^n - u_j^n}{h} = 0, & \text{若 } a_j < 0, \end{cases}$$

其中 $a_j = a(x_j)$.

这是变系数方程, 为了导出稳定性条件, 用局部固定系数法或视变系数为常系数法, 即把 a_j 看成与下标无关, 再用 Fourier 方法, 同样得到条件 (4.3.4) 及 (4.3.5), 只不过用 a_j 代替那里的 a 就是了.

将 (4.3.4),(4.3.5) 统一写成

$$(4.3.7) \qquad \frac{\tau}{h}\max_j |a_j| \leqslant 1.$$

不难证明,(4.3.7) 是 (4.3.6) 稳定的充分条件. 实际上, 将 (4.3.6) 写成形式 $(4.3.2)_1'$, $(4.3.2)_2'$, 其中 $r = (\tau/h)a_j$. 由条件 (4.3.7), 知 $(4.3.2)_1'$, $(4.3.2)_2'$ 右端系数非负. 当 $a_j \geqslant 0$ 时,

$$|u_j^{n+1}| \leqslant r|u_{j-1}^n| + (1-r)|u_j^n| \leqslant \|\boldsymbol{U}^n\|_\infty,$$

当 $a_j \leqslant 0$ 时,

$$|u_j^{n+1}| \leqslant (1+r)|u_j^n| + (-r)|u_{j+1}^n| \leqslant \|\boldsymbol{U}^n\|_\infty,$$

其中 $\boldsymbol{U}^n$ 是以 u_j^n 为分量的向量. 总之, $\|\boldsymbol{U}^{n+1}\|_\infty \leqslant \|\boldsymbol{U}^n\|_\infty$. 这说明 (4.3.6) 稳定.

按照气体力学的含义 ($a(x)$ 表示气流速度), 称 (4.3.6) 为**迎风格式** (Upwind Scheme). 迎风格式也可推广到线性双曲方程组.

设有线性双曲型方程组:

$$\frac{\partial \boldsymbol{u}}{\partial t}+\boldsymbol{A}\frac{\partial \boldsymbol{u}}{\partial x}=\boldsymbol{f}. \tag{4.3.8}$$

设 $\lambda_1<\lambda_2<\cdots<\lambda_m$ 是矩阵 $\boldsymbol{A}$ 的特征值, $\boldsymbol{l}^{(1)},\cdots,\boldsymbol{l}^{(m)}$ 是相应的左特征向量. 像 §2 那样, 用 $\boldsymbol{l}^{(i)}$ 左乘 (4.3.8), 则得特征关系:

$$\sum_{k=1}^{m} l_k^{(i)}\left(\frac{\partial u_k}{\partial t}+\lambda_i\frac{\partial u_k}{\partial x}\right)=\sum_{k=1}^{m} l_k^{(i)} f_k,\quad i=1,2,\cdots,m. \tag{4.3.9}$$

记 $\lambda_{ij}=\lambda_i(x_j)$,

$$\lambda_{ij}\Delta^* u_{kj}^n=\begin{cases}\lambda_{ij}\dfrac{u_{kj}^n-u_{k,j-1}^n}{h}, & 若\ \lambda_{ij}\geqslant 0,\\[2ex] \lambda_{ij}\dfrac{u_{k,j+1}^n-u_{kj}^n}{h}, & 若\ \lambda_{ij}<0,\end{cases} \tag{4.3.10}$$

则逼近 (4.3.9) 的迎风差分格式是

$$\sum_{k=1}^{m} l_{kj}^{(i)}\left(\frac{u_{kj}^{n+1}-u_{kj}^n}{\tau}+\lambda_{ij}\Delta^* u_{kj}^n\right)=\sum_{k=1}^{m} l_{kj}^{(i)} f_{kj}, \tag{4.3.11}$$

其中 $i=1,2,\cdots,m$, $l_{kj}^{(i)}=l_k^{(i)}(x_j)$. 稳定性条件仍然是

$$|\tau\lambda_{ij}/h|\leqslant 1,\quad i=1,2,\cdots,m,\quad j=0,\pm 1,\cdots. \tag{4.3.12}$$

格式 (4.3.11) 是 Courant, Isaacson 和 Rees(1952) 提出的, 并证明了在条件 (4.3.12) 下差分解对光滑解的收敛性.

3.2 积分守恒差分格式

前面介绍的迎风格式是根据特征走向造出的向前或向后差分格式. 现在从积分守恒方程出发构造差分格式.

所谓守恒方程是指如下散度型偏微分方程:

$$\frac{\partial u}{\partial t}+\frac{\partial f(x,u)}{\partial x}=0. \tag{4.3.13}$$

设 G 是 xt 平面任一有界域, 据 Green 公式,

$$\iint\limits_G\left(\frac{\partial u}{\partial t}+\frac{\partial f}{\partial x}\right)\mathrm{d}x\mathrm{d}t=\int_\Gamma (f\mathrm{d}t-u\mathrm{d}x),$$

其中 $\Gamma = \partial G$ (取逆时针方向). 于是可将 (4.3.13) 写成积分守恒形式

$$\int_{\Gamma} (f\mathrm{d}t - u\mathrm{d}x) = 0. \tag{4.3.14}$$

我们先从 (4.3.14) 出发构造熟知的 Lax–Friedrichs 格式. 设网格如图 3.2, 取 G 为以 $A(j+1,n), B(j+1,n+1), C(j-1,n+1)$ 和 $D(j-1,n)$ 为顶点的开矩形,$\Gamma = \overline{ABCDA}$ 为其边界 (取逆时针方向), 则

$$\int_{\Gamma} (f\mathrm{d}t - u\mathrm{d}x) = \int_{DA} (-u)\mathrm{d}x + \int_{BC} (-u)\mathrm{d}x + \int_{AB} f\mathrm{d}t + \int_{CD} f\mathrm{d}t. \tag{4.3.15}$$

右端第一个积分用梯形公式, 第二个积分用中矩形公式, 第三、四两个积分用下矩形公式, 则由 (4.3.14), (4.3.15) 得 Lax–Friedrichs 格式:

$$\frac{u_j^{n+1} - \frac{1}{2}(u_{j-1}^n + u_{j+1}^n)}{\tau} + \frac{f_{j+1}^n - f_{j-1}^n}{2h} = 0, \tag{4.3.16}$$

其中 $f_j^n = f(x_j, u_j^n)$. Lax–Friedrichs 格式截断误差的阶是 $O(\tau + h^2)$.

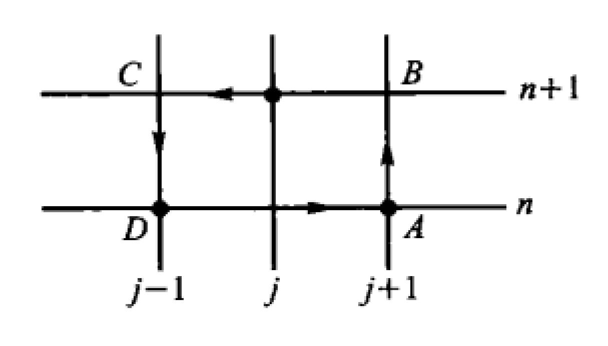


图 3.2

特别当 $f = au$ 时, Lax–Friedrichs 格式相当于 $(4.3.2)_3$ 中的 u_j^n 代以 $\frac{1}{2}(u_{j-1}^n + u_{j+1}^n)$ 的结果. 我们知道, $(4.3.2)_3$ 恒不稳定, 但由 Fourier 方法可知, Lax–Friedrichs 格式稳定的充要条件是

$$\frac{|a|\tau}{h} \leqslant 1. \tag{4.3.17}$$

若在上述推导过程中将 u 换成向量 $\boldsymbol{U} = (u_1, \cdots, u_m)^{\mathrm{T}}, f$ 换成向量 $\boldsymbol{f} = (f_1, \cdots, f_m)^{\mathrm{T}}$, 则得到方程组的 Lax–Friedrichs 格式.

现在由积分守恒方程导出另一种所谓**盒式格式** (Box Scheme). 如图 3.3, 取 G 为以网点 $A(j,n), B(j,n+1), C(j-1,n+1)$ 和 $D(j-1,n)$ 为顶点的矩形, Γ 是 G 的边界. 此时积分型方程仍具有形式 (4.3.15). 右端各项积分用梯形公式近似, 则得

$$\frac{u_j^{n+1} - u_j^n}{\tau} + \frac{u_{j-1}^{n+1} - u_{j-1}^n}{\tau} + \frac{f_j^n - f_{j-1}^n}{h} + \frac{f_j^{n+1} - f_{j-1}^{n+1}}{h} = 0. \tag{4.3.18}$$

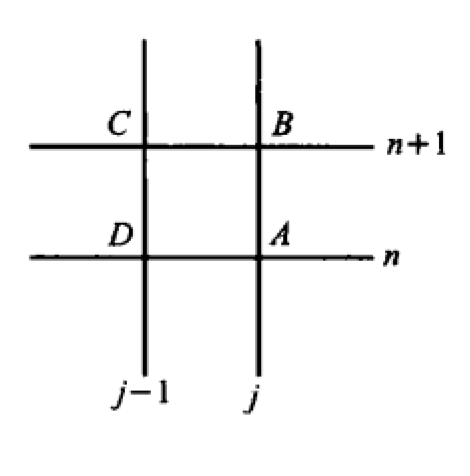


图 3.3

特别以 $f=au$ 代入, 并令 $r=\tau/h$, 得

$$(1+au)u_j^{n+1}+(1-ar)u_{j-1}^{n+1}=(1-ar)u_j^n+(1+ar)u_{j-1}^n. \tag{4.3.19}$$

当 $a>0$ 时, 边值给在左端, 计算由左向右进行; 当 $a<0$ 时, 边值给在右端, 计算由右向左进行, 这些计算都是显的. 用 Fourier 方法还可证明 (4.3.19) 恒稳定. 若 a 依赖 x, 则用 $a_{j-\frac{1}{2}}$ 代替 (4.3.19) 中的 a. 像迎风格式一样, 盒式格式也可用于双曲方程组. 此外, 若视 (4.3.18) 是在 $\left(x_{j-\frac{1}{2}},t_{n+\frac{1}{2}}\right)$ 的差分逼近, 则截断误差的阶为 $O(\tau^2+h^2)$.

3.3 粘性差分格式

粘性差分格式的构造分两步. 先在双曲方程中引进一带二阶空间导数的小参数项, 称为**粘性项**, 使之成为一带小参数的抛物方程, 例如

$$\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=\varepsilon\frac{\partial^2 u}{\partial x^2}\quad(\varepsilon>0); \tag{4.3.20}$$

然后构造逼近相应抛物方程的差分格式. 自然要求 $\varepsilon\to 0$, 当 $\tau\to 0$.

许多逼近双曲方程的差分格式可看作粘性差分格式. 例如迎风格式 (4.3.6) 可改写为

$$\frac{u_j^{n+1}-u_j^n}{\tau}+a_j\frac{u_{j+1}^n-u_{j-1}^n}{2h}=\frac{h}{2}a_j\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2},\quad a_j\geqslant 0,$$

$$\frac{u_j^{n+1}-u_j^n}{\tau}+a_j\frac{u_{j+1}^n-u_{j-1}^n}{2h}=-\frac{h}{2}a_j\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2},\quad a_j<0.$$

或写成统一形式

$$\frac{u_j^{n+1}-u_j^n}{\tau}+a_j\frac{u_{j+1}^n-u_{j-1}^n}{2h}=\frac{h}{2}|a_j|\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2}. \tag{4.3.21}$$

这相当于下列带小参数的抛物方程的中心差分逼近:

$$\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=\frac{h}{2}|a|\frac{\partial^2 u}{\partial x^2}. \tag{4.3.22}$$

对于线性双曲方程组 (4.3.8), 设 $\boldsymbol{A}=\boldsymbol{S\Lambda S}^{-1}, \boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$ 是由 $\boldsymbol{A}$ 的特征值组成的对角矩阵, 则可将逼近 (4.3.8) 的迎风格式写成:

$$\frac{\boldsymbol{u}_j^{n+1}-\boldsymbol{u}_j^n}{\tau}+\boldsymbol{A}\frac{\boldsymbol{u}_{j+1}^n-\boldsymbol{u}_{j-1}^n}{2h}=\frac{h}{2}\boldsymbol{S}\left|\boldsymbol{\Lambda}\right|\boldsymbol{S}^{-1}\frac{\boldsymbol{u}_{j+1}^n-2\boldsymbol{u}_j^n+\boldsymbol{u}_{j-1}^n}{h^2} \tag{4.3.23}$$

其中 $|\boldsymbol{\Lambda}|=\mathrm{diag}(|\lambda_1|,\cdots,|\lambda_n|)$.

Lax–Friedrichs 格式 (4.3.16) 也可改写成:

$$\frac{u_j^{n+1}-u_j^n}{\tau}+\frac{f_{j+1}^n-f_{j-1}^n}{2h}=\frac{h^2}{2\tau}\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2}. \tag{4.3.24}$$

可看成是带小参数的抛物方程:

$$\frac{\partial u}{\partial t}+\frac{\partial f}{\partial x}=\frac{h}{2r}\frac{\partial^2 u}{\partial x^2} \tag{4.3.25}$$

中心差分化的结果, 网比 $r=\tau/h$ 固定.

比较 (4.3.22) 和 (4.3.25) 可知, 小参数的不同取法可导出各种不同的差分格式, 所以选取小参数是构造粘性差分格式的关键. 对于实际问题, 可以直接取自然粘性项, 也可以根据物理上的某些考虑构造人工粘性项. von Neumann 和 Richtmyer 引进的粘性项就属于后一种 (参看 [28] §12.10). 现在以 Lax–Wendroff 格式为例, 介绍另一种引进人工粘性项的方法.

为简便起见, 不妨设 (4.3.13) 中的 $f=f(u)$ (不显含 x). 将 u 关于时间变量 t 展开, 有

$$u(x_j,t_{n+1})=u(x_j,t_n)+\tau\left(\frac{\partial u}{\partial t}\right)_j^n+\frac{\tau^2}{2}\left(\frac{\partial^2 u}{\partial t^2}\right)_j^n+\cdots.$$

利用 (4.3.13) 将关于 t 的偏导数换成关于 x 的偏导数, 并注意

$$\frac{\partial^2 u}{\partial t^2}=-\frac{\partial}{\partial t}\frac{\partial f}{\partial x}=-\frac{\partial}{\partial x}\frac{\partial f}{\partial t}=-\frac{\partial}{\partial x}\left(f'(u)\frac{\partial u}{\partial t}\right)=\frac{\partial}{\partial x}\left(f'(u)\frac{\partial f}{\partial x}\right),$$

则

$$u(x_j,t_{n+1})=u(x_j,t_n)-\tau\left(\frac{\partial f}{\partial x}\right)_j^n+\frac{\tau^2}{2}\left(\frac{\partial}{\partial x}\right)\left(f'(u)\frac{\partial f}{\partial x}\right)_j^n+\cdots.$$

然后略去余项, 并用中心差商代替对 x 的偏导数, 则得 Lax–Wendroff 格式:

$$\begin{aligned}u_j^{n+1}=u_j^n-\frac{1}{2}\frac{\tau}{h}(f_{j+1}^n-f_{j-1}^n)+\\ \frac{1}{2}\left(\frac{\tau}{h}\right)^2[a_{j+\frac{1}{2}}^n(f_{j+1}^n-f_j^n)-a_{j-\frac{1}{2}}^n(f_j^n-f_{j-1}^n)]\end{aligned} \tag{4.3.26}$$

其中 $a_{j+\frac{1}{2}}^n=f'\left(\frac{1}{2}u_{j+1}^n+\frac{1}{2}u_j^n\right)$. 特别, 当 a 是常数时,(4.3.26) 为

$$u_j^{n+1}=u_j^n-\frac{1}{2}a\frac{\tau}{h}(u_{j+1}^n-u_{j-1}^n)+\frac{1}{2}\left(a\frac{\tau}{h}\right)^2(u_{j+1}^n-2u_j^n+u_{j-1}^n). \tag{4.3.27}$$

Lax–Wendroff 格式可看成是带粘性项方程

$$\frac{\partial u}{\partial t}+\frac{\partial f}{\partial x}=\frac{\tau}{2}\frac{\partial}{\partial x}\left(f'(u)\frac{\partial f}{\partial x}\right) \tag{4.3.28}$$

的中心差分格式. 显然, 格式 (4.3.26) 的截断误差的阶是 $O(\tau^2+h^2)$. 由 Fourier 方法可知 (4.3.27) 的增长因子是

$$G=1-\mathrm{i}ra\sin(\alpha h)-(ra)^2(1-\cos(\alpha h)),$$

其中 $r=\tau/h$. 为使 $|G|\leqslant 1$, 必须且只须 $r|a|\leqslant 1$. 这就是稳定性条件.

为了避免出现导数 $a=f'(u)$, 在网格中心引进过渡值

$$u_{j+\frac12}^{n+\frac12}=\frac12(u_{j+1}^n+u_j^n)-\frac{r}{2}(f_{j+1}^n-f_j^n), \tag{4.3.29$_1$}$$

再由下式得到最终值:

$$u_j^{n+1}=u_j^n-r\left(f_{j+\frac12}^{n+\frac12}-f_{j-\frac12}^{n+\frac12}\right). \tag{4.3.29$_2$}$$

称 $(4.3.29)_1$ 和 $(4.3.29)_2$ 为两步 Lax–Wendroff 法, 它仍然有二阶截断误差, 且当 $a=f'(x)=$ 常数时可转化成 (4.3.27).

粘性差分格式 (4.3.24) 和 (4.3.26) 也可直接推广到双曲方程组, 只需将那里的 u 和 f 换成 m 维向量函数.

3.4 其他差分格式

下面列出逼近 (4.3.1) 的其他一些差分格式及相关结果, 这些格式各有其特点. 以下用 $r=a\tau/h$ 表示网比, 并令

$$\Delta_+u_j=u_{j+1}-u_j,\Delta_-u_j=u_j-u_{j-1},\Delta_0u_j=u_{j+1}-u_{j-1}.$$

1. Beam-Warming 格式:

$$\begin{cases}u_j^*=u_j^n-r\Delta_-u_j^n,\\ u_j^{n+1}=\dfrac12\left(u_j^n+u_j^*-r\Delta_-u_j^*-r\Delta_-\Delta_+u_{j-1}^n\right).\end{cases} \tag{4.3.30}$$

稳定性条件: $0\leqslant|r|\leqslant 2$, 截断误差阶: $O\left(\tau^2\right)+O\left(\tau h\right)+O\left(h^2\right)$.

2. MacCormack 格式:

$$\begin{cases}u_j^*=u_j^n-r\Delta_+u_j^n,\\ u_j^{n+1}=\dfrac12\left(u_j^n+u_j^*-r\Delta_-u_j^*\right).\end{cases} \tag{4.3.31}$$

稳定性条件: $|r|\leqslant 1$. 截断误差阶: $O\left(\tau^2\right)+O\left(h^2\right)$.

3. 隐式迎风格式:

$$
(4.3.32)\qquad \begin{cases} \dfrac{u_j^{n+1}-u_j^n}{\tau}+a\dfrac{u_j^{n+1}-u_{j-1}^{n+1}}{h}=0, & 若\ a\geqslant 0,\\ \dfrac{u_j^{n+1}-u_j^n}{\tau}+a\dfrac{u_{j+1}^{n+1}-u_j^{n+1}}{h}=0, & 若\ a<0, \end{cases}
$$

恒稳定. 截断误差阶: $O(\tau)+O(h)$.

4. 隐式中心格式:

$$
(4.3.33)\qquad \frac{u_j^{n+1}-u_j^n}{\tau}+a\frac{u_{j+1}^{n+1}-u_{j-1}^{n+1}}{2h}=0.
$$

恒稳定. 截断误差阶: $O(\tau)+O(h^2)$.

5. 跳蛙 (Leap-frog) 格式:

$$
(4.3.34)\qquad \frac{u_j^{n+1}-u_j^{n-1}}{2\tau}+a\frac{u_{j+1}^n-u_{j-1}^n}{2h}=0.
$$

稳定条件: $|a\tau/h|\leqslant 1$. 截断误差阶: $O(\tau^2)+O(h^2)$.

习　题

1. 逼近方程

$$
\frac{\partial u}{\partial t}+\frac{\partial(au)}{\partial x}=0\quad (a=a(x))
$$

的另一形式的迎风格式为

$$
(4.3.35)_1\qquad \frac{u_j^{n+1}-u_j^n}{\tau}+\frac{a_{j+\frac12}u_{j+\frac12}^n-a_{j-\frac12}u_{j-\frac12}^n}{h}=0,
$$

其中

$$
(4.3.35)_2\qquad u_{j+\frac12}^n=\begin{cases} u_{j+1}^n, & 若\ a_{j+\frac12}<0,\\ u_j^n, & 若\ a_{j+\frac12}>0,\end{cases}
$$

$$
(4.3.35)_3\qquad u_{j-\frac12}^n=\begin{cases} u_j^n, & 若\ a_{j-\frac12}<0,\\ u_{j-1}^n, & 若\ a_{j-\frac12}>0.\end{cases}
$$

因 a 表示流速, 所以也称 $(4.3.35)_{1-2-3}$ 为**偏上游格式** (Upstream Scheme). 证明当 $\tau\leqslant h/\sup|a(x)|$ 时格式稳定.

2. 利用特征构造逼近方程 (4.3.1) 的两种隐式迎风格式. 第一种, 如图 3.4(a), 过 P_0 作特征, 若特征与对角线 $\overline{Q_0P_{-1}}$ 相交, 则利用 $u_{Q_0},u_{P_{-1}}$ 作交点的线性插值, 并取 u_{P_0} 等于交点值. 若特征与对角线 $\overline{Q_0P_1}$ 相交, 可类似的作 u_{P_0}. 第二种, 如图 3.4(b), 过 P_0 作特征, 则特征必与线段 $\overline{Q_0Q_1},\overline{Q_1P_1},\overline{Q_0Q_{-1}}$ 和 $\overline{Q_{-1}P_{-1}}$ 之一相交, 例如和 $\overline{Q_1P_1}$ 相交, 则利用 u_{Q_1},u_{P_1} 作交点的线性插值, 并取 u_{P_0} 等于交点值, 余类推. 试导出计算公式并研究稳定性.

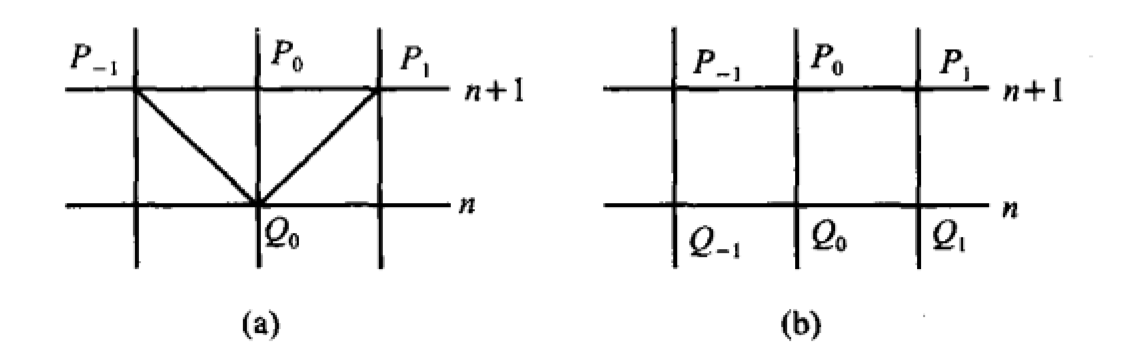


图 3.4

3. (实习题) 利用显和隐的迎风格式计算

$$
\begin{aligned}
&u_t - 2u_x = 0, \quad x \in (0,1), \quad t > 0, \\
&u(x,0) = 1 + \sin 2\pi x, \quad x \in [0,1], \\
&u(1,t) = 1.
\end{aligned}
$$

取 $h = 0.1, \tau = 0.02$, 算出解在 $t = 0.1, 0.5$ 的值.

4. (实习题) 求解初边值问题:

$$
\begin{cases} \dfrac{\partial u}{\partial t} + \dfrac{\partial u}{\partial x} = 0, \quad 0 < x < \infty, \quad t > 0, \\ u(x,0) = |x-1|, \quad u(0,t) = 1. \end{cases} \tag{4.3.36}
$$

(1) 利用格式 $(4.3.2)_1$, 取 $\tau = h = 0.5$, 计算 $t = 1,2,3,4,5$ 的解.

(2) 利用格式 (4.3.32), 取 $\tau = 1, h = 0.5$, 计算 $t = 1,2,3,4,5$ 的解.

注 方程 (4.3.36) 的特征族为: $x - t = c$. 据初边值条件精确解为

$$
u(x,t) = \begin{cases} 1, \quad -\infty < x - t < 0, \\ 1 - (x-t), \quad 0 < x - t < 1, \\ (x-t) - 1, \quad 1 < x - t < \infty. \end{cases}
$$

§4 初边值问题和对流占优扩散方程

4.1 初边值问题

以模型问题

$$
\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0, \quad x \in (0,1), \quad t > 0, \tag{4.4.1}
$$

$$
u(x,0) = \varphi(x), \quad x \in [0,1] \tag{4.4.2}
$$

为例介绍边值条件几种给法及其逼近方法.

1. **周期边值条件** 是指边值条件:

$$u(0,t)=u(1,t).$$

设初值函数也以 1 为周期: $\varphi(0)=\varphi(1)$. 则可将此初边值问题以周期 1 扩展到 x- 轴, 使其成为纯初值问题. 设 $a>0, r=a\tau/h$, 则求解它的迎风格式为

$$u_j^{n+1}=u_j^n-r(u_j^n-u_{j-1}^n),\quad j=1,2,\cdots,J, \tag{4.4.3}$$
$$u_0^{n+1}=u_J^{n+1},$$
$$u_j^0=\varphi_j=\varphi(x_j),\quad j=1,2,\cdots,J, \tag{4.4.4}$$
$$n=0,1,\cdots$$

稳定性条件为 $r\leqslant 1$.

若用 Lax–Friedrichs 格式, 则

$$u_j^{n+1}=\frac{1}{2}(1+r)u_{j+1}^n+\frac{1}{2}(1-r)u_{j-1}^n,\quad j=1,2,\cdots,J-1, \tag{4.4.5}$$
$$u_0^{n+1}=u_J^{n+1}, n=0,1,\cdots,$$
$$u_j^0=\varphi_j,\quad j=1,2,\cdots,J-1.$$

因 $u_{-1}^n=u_{J-1}^n$, 所以

$$u_0^{n+1}=\frac{1}{2}(1+r)u_1^n+\frac{1}{2}(1-r)u_{J-1}^n.$$

2. **Dirichlet 条件** 此时要按照 a 的符号配置边值. 根据特征的走向, 边值应如下配置:

$$a>0\text{ 时},\quad u(0,t)=u_0(x),$$
$$a<0\text{ 时},\quad u(1,t)=u_1(x).$$

迎风格式应为

$$u_j^{n+1}=u_j^n-r(u_j^n-u_{j-1}^n),\ \text{当 } a>0,$$
$$u_j^{n+1}=u_j^n-r(u_{j+1}^n-u_j^n),\ \text{当 } a<0.$$

另一个合理的选择是采用隐式迎风格式 (见 §3 (4.3.32)):

$$\begin{cases}\dfrac{u_j^{n+1}-u_j^n}{\tau}+a\dfrac{u_j^{n+1}-u_{j-1}^{n+1}}{h}=0,\quad a>0,\\[2mm] \dfrac{u_j^{n+1}-u_j^n}{\tau}+a\dfrac{u_{j+1}^{n+1}-u_j^{n+1}}{h}=0,\quad a<0.\end{cases}$$

此格式恒稳定, 截断误差的阶为 $O(\tau+h)$. 虽是隐格式, 但可显式求解.

3. **数值边界条件** 若用空间中心差分格式逼近 (4.4.1) 的 Dirichlet 边值问题, 例如用 Lax–Friedrichs 格式或 Lax–Wendrof 格式, 就会发现还缺少一个边值条件, 因此需适当补充一个条件, 称为数值边界条件. 例如用 Lax-Wendrof 格式

$$u_j^{n+1}=u_j^n-\frac{1}{2}r(u_{j+1}^n-u_{j-1}^n)+\frac{1}{2}r^2(u_{j+1}^n-2u_j^n+u_{j-1}^n) \tag{4.4.6}$$

逼近 (4.4.1), 其中 $a<0, x\in(0,1)$, 边值条件为 $u(1,t)=1$. 此时应在 $x=0$ 补充给一边值条件. 通常可按 (偏右) 迎风格式给为

$$u_0^{n+1}=u_0^n-r(u_1^n-u_0^n) \tag{4.4.7}$$

或

$$u_0^{n+1}=u_0^n-r(u_1^{n+1}-u_0^{n+1}).$$

也可用外推数值边值条件:

$$u_0^{n+1}-u_1^{n+1}=0 \tag*{(4.4.8)$_1$}$$

或

$$u_0^{n+1}-2u_1^{n+1}+u_2^{n+1}=0 \tag*{(4.4.8)$_2$}$$

等, 但不能任意指定 u_0^n 值 (参看 [31] 的第 8 章).

4.2 对流占优扩散方程

所谓对流占优扩散方程, 是指带对流项的抛物方程:

$$\frac{\partial u}{\partial t}+b\frac{\partial u}{\partial x}=a\frac{\partial^2 u}{\partial x^2}, \tag{4.4.9}$$

其中 a,b 是常数, $a>0$, 而 $a\ll|b|$ (即 $|b|$ 相对 a 充分大). 此时 (4.4.9) 虽是抛物方程, 但其解却具双曲性质, 构造差分格式时需考虑解的这一性质. 先处理方程 (4.4.9) 的左端. 令

$$\alpha=(1+b^2)^{\frac{1}{2}}.$$

与 u_t+bu_x 相伴的特征方向为

$$\nu=\left(\frac{1}{\alpha},\frac{b}{\alpha}\right),$$

沿 ν 的方向导数

$$\frac{\partial}{\partial\nu}=\frac{1}{\alpha}\frac{\partial}{\partial t}+\frac{b}{\alpha}\frac{\partial}{\partial x}.$$

所以 (4.4.9) 可写成形式:

$$\alpha\frac{\partial u}{\partial \nu}=a\frac{\partial^2 u}{\partial x^2}. \tag{4.4.10}$$

取时间步长 $\tau>0$, 沿 t 轴取节点 $t=t_n=n\tau,\ n=1,2,\cdots$. 由 (x,t_n) 出发的特征 (方向为 ν) 与直线 $t=t_{n-1}$ 交于

$$\overline{x}=x-b\tau$$

(参看图 4.1). 自然用下式逼近沿特征方向的导数:

$$\alpha\frac{\partial u}{\partial \nu}\approx\alpha\frac{u(x,t_n)-u(\overline{x},t_{n-1})}{[(x-\overline{x})^2+\tau^2]^{\frac{1}{2}}}=\frac{u(x,t_n)-u(\overline{x},t_{n-1})}{\tau}.$$

于是可用如下方程近似替代 (4.4.10):

$$\frac{u(x,t_n)-u(\overline{x},t_{n-1})}{\tau}=a\frac{\partial^2 u}{\partial x^2}. \tag{4.4.11}$$

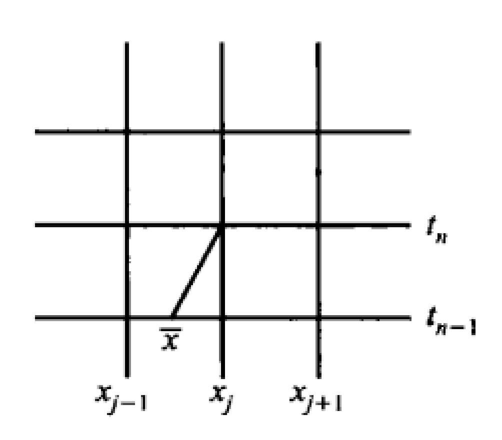


图 4.1

取 $x=x_j,\tau>0$ 充分小, 使 $(\overline{x},t_{n-1})$ 位于 $(x_{j-1},t_{n-1}),(x_{j+1},t_{n-1})$ 之间, 则 $b\geqslant 0$ 时位于 (x_j,t_{n-1}) 左侧, $b<0$ 时位于 (x_j,t_{n-1}) 右侧 (见图 4.1). 今设 $b\geqslant 0$, 与节 3.1 类似, 在 (x_{j-1},t_{n-1}), (x_j,t_{n-1}) 之间以 u_{j-1}^{n-1},u_j^{n-1} 为型值作线性插值, 得

$$\begin{aligned}u(\overline{x},t_{n-1})&=\frac{x_j-\overline{x}}{h}u_{j-1}^{n-1}+\frac{\overline{x}-x_{j-1}}{h}u_j^{n-1}\\&=\frac{b\tau}{h}u_{j-1}^{n-1}+\frac{h-b\tau}{h}u_j^{n-1}.\end{aligned} \tag{4.4.12}$$

以之代到 (4.4.11) 左端, 并用二阶中心差商代替右端的二阶偏微商, 就得到逼近 (4.4.9) 的迎风差分格式

$$\frac{u_j^n-u_j^{n-1}}{\tau}+b\frac{u_j^{n-1}-u_{j-1}^{n-1}}{h}=a\frac{u_{j+1}^{n-1}-2u_j^{n-1}+u_{j-1}^{n-1}}{h^2}. \tag{4.4.13}$$

同样, 当 $b<0$ 时有逼近 (4.4.9) 的迎风差分格式:

$$\frac{u_j^n-u_j^{n-1}}{\tau}+b\frac{u_{j+1}^{n-1}-u_j^{n-1}}{h}=a\frac{u_{j+1}^{n-1}-2u_j^{n-1}+u_{j-1}^{n-1}}{h^2}. \tag{4.4.14}$$

若 $a>|b|$, 则用中心差分格式:

$$\frac{u_j^n-u_j^{n-1}}{\tau}+b\frac{u_{j+1}^{n-1}-u_{j-1}^{n-1}}{h}=a\frac{u_{j+1}^{n-1}-2u_j^{n-1}+u_{j-1}^{n-1}}{h^2} \tag{4.4.15}$$

若 $|b|\gg a$ 时仍用 (4.4.15), 则可能出现不应有的震荡. 由第三章节 7.3 知道, (4.4.15) 稳定的充要条件是 (见 (3.7.7))

$$\frac{1}{2}\left(\frac{b\tau}{h}\right)^2\leqslant\frac{a\tau}{h^2}\leqslant\frac{1}{2} \tag{4.4.16}$$

或

$$\frac{1}{2}r_1^2\leqslant r\leqslant\frac{1}{2},$$

其中

$$r_1=\frac{b\tau}{h},r=\frac{a\tau}{h^2}.$$

由 (4.4.16) 得 $\tau\leqslant 2a/|b|$, 若 $|b|\gg a$, 这条件很难满足.

差分格式 (4.4.13), (4.4.14) 的截断误差的阶为 $O(\tau+h)$. 如果用 (x_{j-1},t_{n-1}), $(x_j,t_{n-1}),(x_{j+1},t_{n-1})$ 和型值 $u_{j-1}^{n-1},u_j^{n-1},u_{j+1}^{n-1}$ 的二次插值代替线性插值 (4.4.12), 就可得到截断误差为 $O(\tau+h^2)$ 的格式.

4.3 数值例子

求解对流占优扩散方程 (参看第三章节 7.3)

$$u_t+bu_x=au_{xx},\quad t>0,\quad 0<x<\infty, \tag{4.4.17$_1$}$$

其中 $a,b>0$ 是常数. 定解条件为

$$u(x,0)=0,\quad x>0, \tag{4.4.17$_2$}$$

$$u(0,t)=u_0,\quad t\geqslant 0,\quad u_0\text{ 为正常数} \tag{4.4.17$_3$}$$

$$u(\infty,t)=0,\quad t\geqslant 0. \tag{4.4.17$_4$}$$

精确解

$$u(x,t)=\frac{u_0}{2}\left\{\operatorname{erfc}\left(\frac{x-bt}{2\sqrt{at}}\right)+\exp\left(\frac{bx}{a}\right)\operatorname{erfc}\left(\frac{x+bt}{2\sqrt{at}}\right)\right\},\ \operatorname{erfc}(x)=\frac{2}{\sqrt{\pi}}\int_x^\infty \mathrm{e}^{-t^2}\mathrm{d}t.$$

现在采用隐式迎风格式 (参看 (4.4.13)):

$$\frac{u_j^{n+1}-u_j^n}{\tau}+b\frac{u_j^{n+1}-u_{j-1}^{n+1}}{h}=a\frac{u_{j+1}^{n+1}-2u_j^{n+1}+u_{j-1}^{n+1}}{h^2}$$

(i) 设 $a=1$, $b=1$. 取 $h=0.5$, $\tau=0.25$. 此时 $r=\tau/h^2=1$. 记 $t_n=n\tau$, 计算 $t=t_1,\cdots,t_{50}$ 的差分解, 没有发生震荡, $t=t_{50}$ 的曲线如图 4.2, 实黑线为精确解, 虚线为近似解.

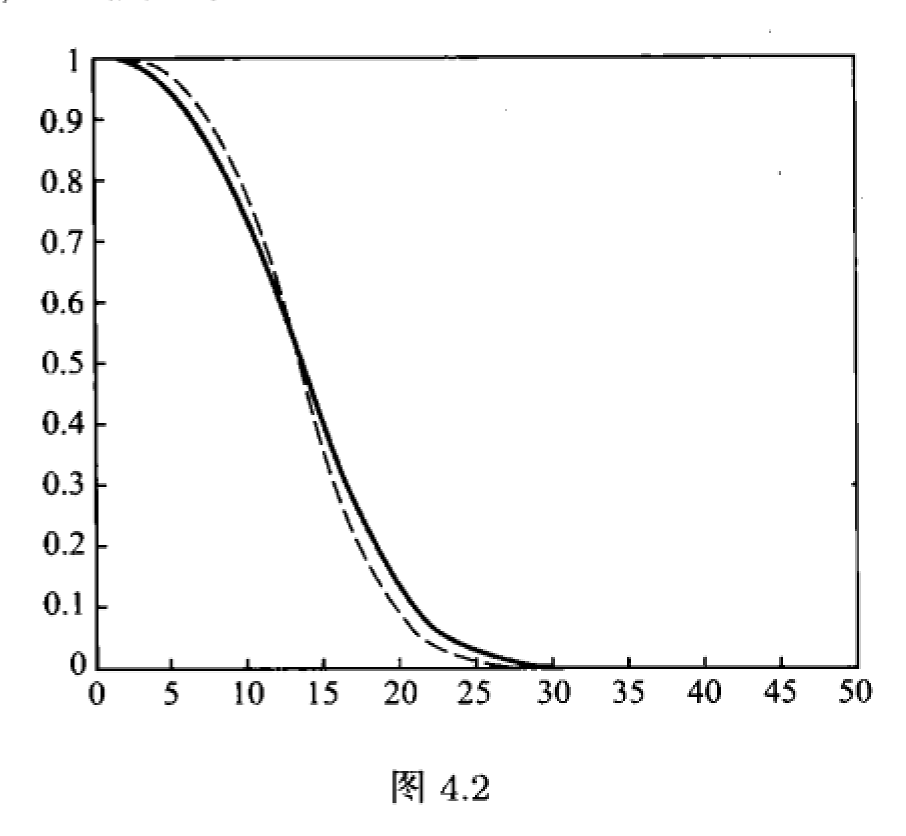


图 4.2

(ii) 设 $a=1, b=10$. 取 $h=0.5$, $\tau=0.25$. 此时 $r=1$. 记 $t_n=n\tau$. 计算 $t=t_1,\cdots,t_{50}$ 的差分解, 没有发生震荡, $t=t_{50}$ 的曲线如图 4.3, 实黑线为精确解, 虚线为近似解. 可见迎风格式计算含大对流项方程仍稳定. 在第三章节 7.3, 我们曾用稳定的显中心差分格式求解问题 (4.4.17), 当对流项 $b=10$ 时解出现不应有的震荡.

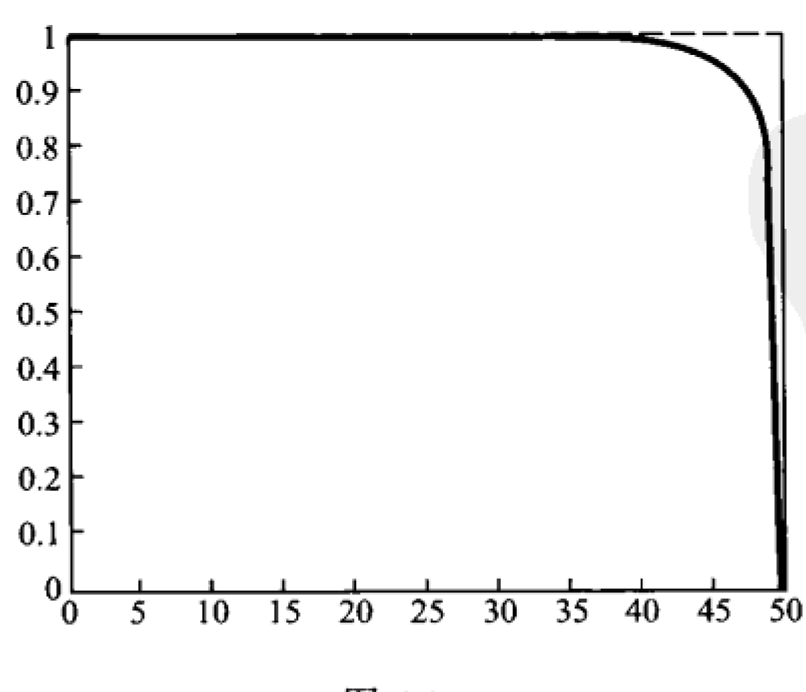


图 4.3

习　　题

1. (实习题) 求解

$$\begin{aligned}&u_t - u_x = 0, \quad x \in (0,1),\\&u(x,0) = \sin^{40} \pi x, \quad x \in [0,1],\\&u(0,t) = u(1,t), \quad t \geqslant 0.\\&(\text{精确解 } u(x,t) = \sin^{40} \pi(x+t))\end{aligned}$$

用迎风格式计算.

(a) 取 $h = 0.05, \tau = 0.04$, 算出 $t = 0.00,\ 0.12,\ 0.20,\ 0.80$ 的值.

(b) 画出 (a) 的图形, 观察峰值位置的变化, 与精确解峰值位置比较.

2. (实习题) 用 Lax–Wendroff 格式求解方程

$$\begin{aligned}&u_t - 2u_x = 0, \quad x \in (0,1), \quad t > 0,\\&u(x,0) = 1 + \sin 2\pi x, \quad x \in [0,1],\\&u(1,t) = 1 + \sin 4\pi t.\end{aligned}$$

(精确解 $u = 1 + \sin 2\pi\,(x + 2t)$)

数值边值条件分别为

(a) $u_0^{n+1} = u_0^n + \dfrac{2\tau}{h}\left(u_1^n - u_0^n\right)$,

(b) $u_0^n = u_1^n$,

(c) $u_0^{n+1} - 2u_1^{n+1} + u_2^{n+1} = 0$,

并与精确解比较.

3. (实习题) 用差分格式

$$\frac{u_j^n - u_j^{n-1}}{\tau} + a\frac{u_{j+1}^{n-1} - u_{j-1}^{n-1}}{h} = \frac{u_{j+1}^{n-1} - 2u_j^{n-1} + u_{j-1}^{n-1}}{h^2}$$

和

$$\frac{u_j^n - u_j^{n-1}}{\tau} + a\frac{u_j^{n-1} - u_{j-1}^{n-1}}{h} = \frac{u_{j+1}^{n-1} - 2u_j^{n-1} + u_{j-1}^{n-1}}{h^2}$$

求解

$$\frac{\partial u}{\partial t} + b\frac{\partial u}{\partial x} = \frac{\partial^2 u}{\partial x^2}$$

初值条件:

$$u(x,0) = \begin{cases} x, & 0 \leqslant x \leqslant \dfrac{1}{2},\\ (1-x), & \dfrac{1}{2} \leqslant x \leqslant 1.\end{cases}$$

边值条件:

$$u(0,t) = u(1,t) = 0.$$

取网比 $r = \frac{1}{2}$, 试就 $b = 1,\ 5, 10,\ 20$ 和步长 $h = 0.1,\ 0.01$, 计算 $t = 0.1,\ 0.5$ 的差分解.



第五章　边值问题的变分形式与 Ritz–Galerkin 法

§1　二次函数的极值

数学物理中的变分原理, 有重要的理论和实际意义, 也是构造微分方程数值解法的基础. 为了便于读者理解, 本节以二次函数的极值问题为例, 介绍变分原理的基本概念和方法. 主要结论和证明思想均可平行推广到线性对称微分方程边值问题.

在 n 维欧氏空间 $\mathbf{R}^n$ 中引进向量、矩阵记号:

$$\boldsymbol{x}=(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}},$$
$$\boldsymbol{b}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}},$$
$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

这里以及往后仍用 $(\ )^{\mathrm{T}}$ 表示括号内向量或矩阵的转置. 令 $\boldsymbol{y}=(\eta_1,\eta_2,\cdots,\eta_n)^{\mathrm{T}}$, 定义 $\boldsymbol{x},\boldsymbol{y}$ 的内积为

$$(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{n}\xi_i\eta_i.$$

考虑 n 个变量的二次函数:

$$F(\boldsymbol{x})=F(\xi_1,\xi_2,\cdots,\xi_n)=\sum_{i,j=1}^{n}\alpha_{ij}\xi_i\xi_j-\sum_{i=1}^{n}b_i\xi_i=(\boldsymbol{A}\boldsymbol{x},\boldsymbol{x})-(\boldsymbol{b},\boldsymbol{x}).$$

它在 $\boldsymbol{x}_0=(\xi_1^{(0)},\xi_2^{(0)},\cdots,\xi_n^{(0)})^{\mathrm{T}}$ 取极值的必要条件是

$$\frac{\partial F(\xi_1^{(0)},\xi_2^{(0)},\cdots,\xi_n^{(0)})}{\partial\xi_k}=\sum_{i=1}^{n}(\alpha_{ik}+\alpha_{ki})\xi_i^{(0)}-b_k=0,\quad k=1,2,\cdots,n.$$

假定 $\alpha_{ik}=\alpha_{ki}$, 即 $\boldsymbol{A}$ 为对称矩阵, 则

$$2\sum_{i=1}^{n}\alpha_{ki}\xi_i^{(0)}=b_k,\quad k=1,2,\cdots,n.$$

不难看出, 若令

$$J(\boldsymbol{x})=\frac{1}{2}(\boldsymbol{A}\boldsymbol{x},\boldsymbol{x})-(\boldsymbol{b},\boldsymbol{x}),\tag{5.1.1}$$

则二次函数 $J(\boldsymbol{x})$ 于 $\boldsymbol{x}_0$ 取极值的必要条件是: $\boldsymbol{x}_0$ 是线性代数方程组

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}\tag{5.1.2}$$

的解.

为了进一步研究 $J(x)$ 于 $\boldsymbol{x}_0$ 的极值性质, 考虑实变量 λ 的二次函数

$$\varphi(\lambda)=J(\boldsymbol{x}_0+\lambda\boldsymbol{x}),$$

其中 $\boldsymbol{x}$ 是任一 n 维非零向量. 若 $J(x)$ 于 $\boldsymbol{x}_0$ 取极小值, 则对任何 $\lambda\neq0,\varphi(\lambda)=J(\boldsymbol{x}_0+\lambda\boldsymbol{x})>J(\boldsymbol{x}_0)=\varphi(0)$, 即 $\varphi(\lambda)$ 于 $\lambda=0$ 取极小值. 反之, 若 $\varphi(\lambda)$ 于 $\lambda=0$ 取极小值, 则对任何非零向量 $\boldsymbol{x},J(\boldsymbol{x}_0+\boldsymbol{x})=\varphi(1)>\varphi(0)=J(\boldsymbol{x}_0)$, 即 $J(\boldsymbol{x})$ 于 $\boldsymbol{x}_0$ 取极小值. 这样, 我们就把多变量函数 $J(\boldsymbol{x})$ 的极值问题化成单变量函数 $\varphi(\lambda)$ 的极值问题.

现在研究 $J(\boldsymbol{x})$ 存在极小值的充分必要条件. 显然

$$\varphi(\lambda)=J(\boldsymbol{x}_0)+\frac{\lambda}{2}[(\boldsymbol{A}\boldsymbol{x}_0,\boldsymbol{x})+(\boldsymbol{A}\boldsymbol{x},\boldsymbol{x}_0)-2(\boldsymbol{b},\boldsymbol{x})]+\frac{\lambda^2}{2}(\boldsymbol{A}\boldsymbol{x},\boldsymbol{x}).$$

因为 $\boldsymbol{A}$ 是对称矩阵, 故

$$\varphi(\lambda)=J(\boldsymbol{x}_0+\lambda\boldsymbol{x})=J(\boldsymbol{x}_0)+\lambda(\boldsymbol{A}\boldsymbol{x}_0-\boldsymbol{b},\boldsymbol{x})+\frac{\lambda^2}{2}(\boldsymbol{A}\boldsymbol{x},\boldsymbol{x}).\tag{5.1.3}$$

若 $J(\boldsymbol{x})$ 于 $\boldsymbol{x}_0$ 取极小值, 则

$$\varphi'(0)=(\boldsymbol{A}\boldsymbol{x}_0-\boldsymbol{b},\boldsymbol{x})=0,\ \text{对任意}\ \boldsymbol{x}\in\mathbf{R}^n,$$

从而 $\boldsymbol{A}\boldsymbol{x}_0-\boldsymbol{b}=\mathbf{0}$, 这说明 $\boldsymbol{x}_0$ 是 (5.1.2) 的解. 又

$$\varphi''(0)=(\boldsymbol{A}\boldsymbol{x},\boldsymbol{x})>0,\ \text{对任意非零向量}\ \boldsymbol{x}\in\mathbf{R}^n,$$

故 $\boldsymbol{A}$ 必为正定矩阵.

反之, 设 $\boldsymbol{A}$ 是对称正定矩阵, $\boldsymbol{x}_0$ 是方程组 (5.1.2) 的解, 即

$$\boldsymbol{A}\boldsymbol{x}_0-\boldsymbol{b}=\mathbf{0},$$

则由 (5.1.3) 得

$$\varphi(\lambda) = J(\boldsymbol{x}_0) + \frac{\lambda^2}{2}(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}) = \varphi(0) + \frac{\lambda^2}{2}(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}) > \varphi(0), \quad \lambda \neq 0, \quad x \neq 0.$$

这说明 $J(x)$ 于 $\boldsymbol{x}_0$ 取极小值. 于是我们得

定理 1.1 设矩阵 $\boldsymbol{A}$ 对称正定, 则下列两问题等价:

(1) 求 $\boldsymbol{x}_0 \in \mathbf{R}^n$ 使

$$J(\boldsymbol{x}_0) = \min_{\boldsymbol{x} \in \mathbf{R}^n} J(\boldsymbol{x}), \tag{5.1.4}$$

其中 $J(\boldsymbol{x})$ 是由 (5.1.1) 定义的二次函数.

(2) 求下列方程组的解:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}. \tag{5.1.5}$$

$J(\boldsymbol{x})$ 是定义在全空间 $\mathbf{R}^n$ 上的二次函数, 称为 $\mathbf{R}^n$ 上的**二次泛函**或简称**泛函数**. 泛函数 $J(\boldsymbol{x})$ 由两部分组成: 第一部分是二次项 $\frac{1}{2}(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x})$, 它由矩阵 $\boldsymbol{A}$ 决定; 第二部分是一次项 $(\boldsymbol{b}, \boldsymbol{x})$, 它由右端向量 $\boldsymbol{b}$ 决定.

定理 1.1 表明, 在矩阵 $\boldsymbol{A}$ 为对称正定的条件下, 若 $\boldsymbol{x}_0$ 是极值问题 (5.1.4) 的解, 则它也是线性方程组 (5.1.5) 的解; 反之亦然. 因此为了确定并计算 $\boldsymbol{x}_0$, 可采取两种不同途径: 一种是求方程组 (5.1.5) 的解, 另一种是求泛函数 $J(\boldsymbol{x})$ 的极小值. 我们更强调这后一途径. 因为许多数学物理问题, 其直接的数学形式就是求意义更广的 “二次泛函” 的极小值, 只是对解做了某些 “光滑性” 假设之后, 才归结到微分方程. 其次, 即便是熟知的微分方程边值问题, 我们也宁愿把它化为某一 “二次泛函” 的极小值问题, 因为从极值问题出发建立数值解法往往更方便.

习　　题

如果 $\varphi'(0) = 0$, 则称 $\boldsymbol{x}_0$ 是 $J(\boldsymbol{x})$ 的**驻点** (或**稳定点**). 设矩阵 $\boldsymbol{A}$ 对称 (不必正定), 求证 $\boldsymbol{x}_0$ 是 $J(\boldsymbol{x})$ 的驻点的充要条件是: $\boldsymbol{x}_0$ 是方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 的解.

§2 Sobolev 空间初步

2.1 弦的平衡

考察一根长为 l 的弦 (见图 2.1), 其两端固定在点 $A(0,0)$ 和 $B(l,0)$. 当没有外力作用时, 它的位置沿水平方向与 x 轴重合. 设有强度为 $f(x)$ 的外荷载垂直向下作用在弦上, 则弦发生形变. 假定荷载很小, 因而发生的形变也很小. 用

$u = u(x)$ 表示在荷载 $f(x)$ 作用下弦的平衡位置. 根据力的平衡条件, $u(x)$ 满足微分方程

$$-Tu'' = f(x), \quad 0 < x < l \tag{5.2.1}$$

和边值条件

$$u(0) = 0, \quad u(l) = 0, \tag{5.2.2}$$

其中 T 是弦的张力 (假定是常数). 这样, 求弦的平衡位置就归结为解两点边值问题 (5.2.1), (5.2.2).

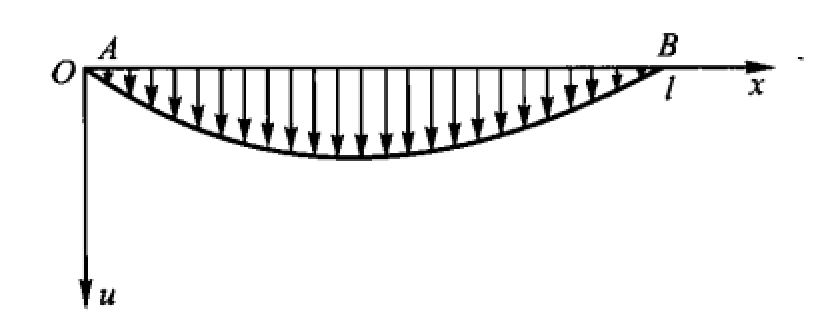


图 2.1

另一方面, 由力学的 "极小位能原理", 弦的平衡位置 (记为 $u_* = u_*(x)$) 是在满足边值条件 (5.2.2) 的一切可能位置中, 使位能取最小者. 设弦处于某一位置 $u = u(x)$, 我们计算它的位能. 此时应变能为

$$W_{内} = \frac{1}{2}\int_0^l T(u')^2 \mathrm{d}x;$$

外力 $f(x)$ 所作的功为

$$W_{外} = -\int_0^l fu\mathrm{d}x,$$

从而总位能为

$$J(u) = W_{内} + W_{外} = \frac{1}{2}\int_0^l [Tu'^2 - 2uf]\mathrm{d}x. \tag{5.2.3}$$

据极小位能原理, $u_* = u_*(x)$ 是如下变分问题的解:

$$J(u_*) = \min_u J(u). \tag{5.2.4}$$

这样, 为了确定弦的平衡位置, 便导致两个不同形式的数学问题: 两点边值问题 (5.2.1), (5.2.2) 和变分问题 (5.2.4). 显然二者之间应具有某种等价关系, 这就是本章要建立的各种变分原理的最简模型.

为了精确地表述变分问题 (5.2.4), 必须指出 $J(u)$ 在哪一个函数类里取极小, 就是说要给出 u 属于哪一个函数空间. 从位能的计算公式 (5.2.3) 看出, 为使积分有意义, 必须对 u, f 作必要的限制. 但又不能限制过严, 以致把取极小值的函数 $u_* = u_*(x)$ 排斥在外. 这样看来, 适当地选取函数空间就十分重要了. 下面我们引进这样的空间.

2.2 一维区间上的 Sobolev 空间 $H^m(I)$

设 $I = (a,b), \overline{I} = [a,b]$. 像通常那样, 用 $L^2(I)$ 表示由定义在 I 上的平方可积的可测函数组成的空间, 其内积和范数分别为

$$(f,g) = \int_a^b fg\mathrm{d}x, \quad f,\ g \in L^2(I),$$

$$||f|| = \sqrt{(f,f)} = \left[\int_a^b |f|^2\mathrm{d}x\right]^{\frac{1}{2}}, \quad f \in L^2(I).$$

我们知道, $L^2(I)$ 关于 "加法" 及 "数乘" 运算是线性空间, 关于 (,) 是完全内积空间, 因此 $L^2(I)$ 是 Hilbert 空间. 所谓完全, 是指 Cauchy 收敛定理在 $L^2(I)$ 中成立. 就是说, $L^2(I)$ 中任一函数列 $\{f_n\}$, 如果关于度量 $||\cdot||$ (即按 $L^2(I)$ 度量) 满足 Cauchy 条件:

$$||f_n - f_m|| \to 0 (n, m \to \infty)$$

(如此的 $\{f_n\}$ 称为基本序列) 则必有 $f \in L^2(I)$ 使得

$$||f_n - f|| \to 0 \quad (n \to \infty),$$

并记为 $\lim\limits_{n\to\infty} f_n = f$. 我们提一下 $L^2(I)$ 空间的一个基本性质: 设 $f, g \in L^2(I)$, 则乘积 $fg \in L^1(I)$, 且

$$|(f,g)| \leqslant ||f|| \cdot ||g||.$$

这是熟知的 Schwarz 不等式 (参看 [2]).

现在看二次泛函 (5.2.3). 假定 $f \in L^2(I)$, 为使能量 $J(u)$ 有意义, 显然应要求 $u \in L^2(I), u' \in L^2(I)$. 因此我们要找的空间应该包含如此的可测函数: u 及其导数 u' 皆属于 $L^2(I)$, 即

$$\int_a^b |u|^2\mathrm{d}x < \infty, \int_a^b |u'|^2\mathrm{d}x < \infty.$$

然而仅限于这样的函数类是不够的, 因为在研究物理、力学等许多自然现象中 (如集中荷载的平板弯曲、具不同介质的弹性体的平衡), 会遇到许多函数并不处处可微; 同时为使要找的空间完全 (即 Cauchy 定理成立), 也需要进一步扩充上述函数类. 为此, 我们引进广义导数 (广义微商) 的概念 (参看 [10]).

用 $C_0^\infty(I)$ 表示于 I 无穷次可微, 且在端点 a,b 的某邻域内 (邻域大小与具体函数有关) 等于零的函数类. 对于任一于 $[a,b]$ 一次连续可微的函数 $f(x)$ 和任意 $\varphi\in C_0^\infty(I)$, 由分部积分法, 恒有

$$\int_a^b f'(x)\varphi(x)\mathrm{d}x=-\int_a^b f(x)\varphi'(x)\mathrm{d}x. \tag{5.2.5}$$

今利用 (5.2.5) 推广导数的概念. 设 $f\in L^2(I)$, 若存在 $g\in L^2(I)$, 使等式

$$\int_a^b g(x)\varphi(x)\mathrm{d}x=-\int_a^b f(x)\varphi'(x)\mathrm{d}x,\quad 对任意\ \varphi\in C_0^\infty(I) \tag{5.2.6}$$

成立, 则说 $f(x)$ 于 I 有**广义导数** $g(x)$, 记为

$$f'(x)=\frac{\mathrm{d}f}{\mathrm{d}x}=g(x).$$

显然, 若 $f(x)$ 在通常意义下有属于 $L^2(I)$ 的导数 $f'(x)$, 则 $f'(x)$ 也是 $f(x)$ 在广义意义下的导数. 相反的结论则不一定成立. 例如 $f(x)=|x|$, 在通常意义下于 $[-1,1]$ 不可微, 但是对任意 $\varphi\in C_0^\infty(I)$,

$$\begin{aligned}-\int_{-1}^1|x|\varphi'(x)\mathrm{d}x&=-\int_0^1 x\varphi'(x)\mathrm{d}x+\int_{-1}^0 x\varphi'(x)\mathrm{d}x\\&=\int_0^1\varphi(x)\mathrm{d}x-\int_{-1}^0\varphi(x)\mathrm{d}x\\&=\int_{-1}^1 g(x)\varphi(x)\mathrm{d}x,\end{aligned}$$

其中

$$g(x)=\begin{cases}-1, & -1\leqslant x<0,\\ 1, & 0\leqslant x\leqslant 1\end{cases}$$

就是 $f(x)$ 的广义导数.

同一个函数 $f(x)$ 的广义导数并不唯一, 但不同的广义导数几乎处处相等. 其证明基于下列

引理 2.1 (变分法的基本引理) 设 $f\in L^2(I)$,

$$\int_a^b f(x)\varphi(x)\mathrm{d}x=0,\ 对任意\ \varphi\in C_0^\infty(I),$$

则 $f(x)$ 几乎处处为 0.

证明 只就 $f\in C(\bar I)$ ($\bar I$ 上的连续函数空间) 证明引理. 这时还可推出更强的结论: $f(x)\equiv 0$. 事实上, 假若 $f(x)$ 不恒等于 0, 不妨设 $f(x)$ 于 $x_0\in(a,b)$ 不

等于 0, 例如 $f(x_0)>0$, 则由连续性, $f(x)$ 必在 x_0 充分小的邻域 $a<x_0-\eta\leqslant x\leqslant x_0+\eta<b$ 内也大于 0. 取

$$\varphi(x)=\begin{cases}\exp\left(-\dfrac{1}{\eta^2-(x-x_0)^2}\right), & |x-x_0|<\eta,\\ 0, & \text{其他},\end{cases}$$

则 $\varphi\in C_0^\infty(I)$, 且

$$\int_a^b f(x)\varphi(x)\mathrm{d}x=\int_{x_0-\eta}^{x_0+\eta}f(x)\exp\left(-\frac{1}{\eta^2-(x-x_0)^2}\right)\mathrm{d}x>0.$$

与假设矛盾.

利用变分法的基本引理可证明 $f(x)$ 的不同的广义导数几乎处处相等.

注意, 并非任何 $f\in L^2(I)$ 都有属于 $L^2(I)$ 的广义导数. 例如, 阶梯函数 $(a=-1,b=1)$

$$f(x)=\begin{cases}0, & -1\leqslant x\leqslant 0,\\ 1, & 0<x<1.\end{cases}$$

若 $f(x)$ 有广义导数 $g(x)$, 则

$$\int_{-1}^1 g(x)\varphi(x)\mathrm{d}x=-\int_{-1}^1 f(x)\varphi'(x)\mathrm{d}x=-\int_0^1\varphi'(x)\mathrm{d}x=\varphi(0), \text{对任意 } \varphi\in C_0^\infty(I).$$

从而 $g(x)=\delta(x)$ —— Dirac 函数, 称为 δ- 函数. 今证 $g(x)$ 不属于 $L^2(I)$. 实际上, 若 $g(x)\in L^2(I)$, 则由 Schwarz 不等式,

$$|\varphi(0)|=\left|\int_{-1}^1 g(x)\varphi(x)\mathrm{d}x\right|\leqslant\|g\|\,\|\varphi\|, \text{对任意 } \varphi\in C_0^\infty(I). \tag{5.2.7}$$

特别对 $0<\varepsilon<1$, 取

$$\varphi(x)=\varphi_\varepsilon(x)=\begin{cases}\exp\left(-\dfrac{1}{1-(x/\varepsilon)^2}\right), & \text{当 } |x|<\varepsilon,\\ 0, & \text{其他},\end{cases}$$

则 $\varphi_\varepsilon(x)\in C_0^\infty(I), \varphi_\varepsilon(0)=\mathrm{e}^{-1}$,

$$\|\varphi_\varepsilon\|=\int_{-\varepsilon}^{\varepsilon}\exp\left(-\frac{2}{1-(x/\varepsilon)^2}\right)\mathrm{d}x=2\varepsilon\int_0^1\exp\left(-\frac{2}{1-t^2}\right)\mathrm{d}t\to 0(\varepsilon\to 0).$$

以此代到 (5.2.7) 并令 $\varepsilon\to 0$, 就导致矛盾.

这个例子告诉我们, δ- 函数不属于 $L^2(I)$, 阶梯函数没有广义导数.

现在定义

$$H^1(I)=\{f|f\in L^2(I),\ f'\in L^2(I)\},$$

其中 f' 是 f 的广义导数. 显然 $H^1(I)$ 是线性空间. 于 $H^1(I)$ 引进内积

$$(f,g)_1 = \int_a^b [fg + f'g']\mathrm{d}x \tag{5.2.8}$$

和范数

$$\|f\|_1 = \sqrt{(f,f)_1} = \left[\int_a^b (f^2 + f'^2)\mathrm{d}x\right]^{\frac{1}{2}}. \tag{5.2.9}$$

可以证明, $H^1(I)$ 是完全内积空间, 因此是 Hilbert 空间, 称之为 **Sobolev 空间**.

同样可定义 m 阶广义导数和 m 阶的 Sobolev 空间 $H^m(I)$, 其内积和范数分别为

$$(f,g)_m = \sum_{k=0}^m \int_a^b f^{(k)}(x)g^{(k)}(x)\mathrm{d}x, \tag{5.2.10}$$

$$\|f\|_m = \sqrt{(f,f)_m} = \left[\sum_{k=0}^m \int_a^b |f^{(k)}(x)|^2\mathrm{d}x\right]^{\frac{1}{2}}. \tag{5.2.11}$$

当 $m=0$ 时, $H^0(I)$ 就是 $L^2(I)$ 空间,

$$(f,g)_0 = (f,g), \quad \|f\|_0 = \|f\|.$$

从位能的表达式 (5.2.3) 看出, 当 $f \in H^0(I), u \in H^1(I)$ 时, 位能 $J(u)$ 恒有意义. 此外, u 还应满足边值条件 (5.2.2). $H^1(I)$ 中所有满足齐次边值条件 (5.2.2) 的函数类构成 $H^1(I)$ 的子空间, 记为 $H_0^1(I)$ 或 H_0^1. H_0^1 就是我们所要找的函数空间. 现在我们可将变分问题 (5.2.4) 精确地叙述为: 求 $u_* \in H_0^1$ 使

$$J(u_*) = \min_{u \in H_0^1} J(u). \tag{5.2.4$'$}$$

注 2.1 前面讨论的是定义在区间上的实值函数, 实际上所有定义和结论都可推广到实变复值函数. 例如设函数 $f = u+\mathrm{i}v$, 其中 u 和 v 为 f 的实部和虚部, 则 f 的 m 阶广义导数

$$f^{(m)} = u^{(m)} + \mathrm{i}v^{(m)},$$

Sobolev 空间

$$H^m = \left\{f \,\middle|\, f^{(s)} \in L^2, s = 0,1,2,\cdots,m\right\},$$

内积为

$$(f,g)_m = \int_a^b \sum_{s=0}^m f^{(s)}\overline{g}^{(s)}\mathrm{d}x,$$

范数

$$\|f\|_m = \sqrt{(f,f)_m}.$$

2.3 平面域上的 Sobolev 空间 $H^m(G)$

节 2.2 介绍的一维区域上的 Sobolev 空间及广义导数等概念, 可以平行地推广到多维区域. 假定 G 是有界平面区域, 其边界 Γ 是按段光滑的简单闭曲线, $\overline{G}=G\bigcup\Gamma$ 是 G 的闭包. 对于 $\overline{G}$ 上的任一函数 $u(x,y)$, 称集合 $\{(x,y)|u(x,y)\neq 0,(x,y)\in\overline{G}\}$ 的闭包为 u 的**支集**. 如果 u 的支集 $\subset G$ 内, 则说 u 于 G 具有**紧致支集**. 显然, 具有紧致支集的函数必在边界 Γ 的某一邻域内恒等于零.

我们用 $C_0^\infty(G)$ 表示 G 上无穷次可微并具有紧致支集的函数类, $L^2(G)$ 是定义在 G 上平方可积的可测函数空间, 其内积和范数分别为

$$(f,g)=\iint\limits_G fg\mathrm{d}x\mathrm{d}y,$$

$$||f||=\sqrt{(f,f)}=\left(\iint\limits_G |f|^2\mathrm{d}x\mathrm{d}y\right)^{\frac{1}{2}}.$$

对 $f\in L^2(G)$, 如果存在 $g,h\in L^2(G)$, 使等式

$$\iint\limits_G g\varphi\mathrm{d}x\mathrm{d}y=-\iint\limits_G f\frac{\partial\varphi}{\partial x}\mathrm{d}x\mathrm{d}y,$$

$$\iint\limits_G h\varphi\mathrm{d}x\mathrm{d}y=-\iint\limits_G f\frac{\partial\varphi}{\partial y}\mathrm{d}x\mathrm{d}y,$$

对任意 $\varphi\in C_0^\infty(G)$ 成立, 则说 f 有对 x 的一阶广义导数 g 和对 y 的一阶广义导数 h, 记作

$$f_x=\frac{\partial f}{\partial x}=g,$$

$$f_y=\frac{\partial f}{\partial y}=h.$$

类似地可定义高阶广义导数.

像一维情形一样, 定义

$$H^1(G)=\{f(x,y)\,|f,f_x\,,f_y\in L^2(G)\},$$

其中 f_x 和 f_y 是 f 的广义导数. 于 $H^1(G)$ 引入内积

$$(f,g)_1=\iint\limits_G [fg+f_xg_x+f_yg_y]\mathrm{d}x\mathrm{d}y \tag{5.2.12}$$

和范数

$$||f||_1=\sqrt{(f,f)}=\left(\iint\limits_G [|f|^2+|f_x|^2+|f_y|^2]\mathrm{d}x\mathrm{d}y\right)^{\frac{1}{2}}, \tag{5.2.13}$$

则 $H^1(G)$ 是 Hilbert 空间, 称之为 **Sobolev 空间**. 还可以定义高阶导数及高阶 Sobolev 空间 $H^m(G)$, m 是非负整数.

往后会看到, 在二阶椭圆型方程的数值解法及理论研究中, $H^1(G)$ 是最常用的函数空间.

习 题

设 $f'(x)$ 为 $f(x)$ 的一阶广义导数, 试用类似办法定义 $f(x)$ 的 k 阶广义导数 $f^{(k)}(x)(k=1,2,\cdots)$.

§3 两点边值问题

3.1 极小位能原理

现在我们考虑一般两点边值问题的变分形式. 为此我们先看 (5.2.3) 中位能 $J(u)$ 的结构. 引进微分算子

$$Lu=-T\frac{\mathrm{d}^2u}{\mathrm{d}x^2},$$

则

$$W_{内}=\frac{1}{2}(Lu,u)=\frac{1}{2}\int_0^l\left(-T\frac{\mathrm{d}^2u}{\mathrm{d}x^2}\right)u\mathrm{d}x=\frac{1}{2}\int_0^l T\left(\frac{\mathrm{d}u}{\mathrm{d}x}\right)^2\mathrm{d}x,$$

$$W_{外}=-(f,u)=-\int_0^l fu\mathrm{d}x.$$

于是

$$J(u)=\frac{1}{2}(Lu,u)-(f,u). \tag{5.3.1}$$

与 (5.1.1) 比较:

$$H_0^1\leftrightarrow\mathbf{R}^n,$$
$$L\leftrightarrow\boldsymbol{A},$$
$$f\leftrightarrow\boldsymbol{b},$$

便知 $J(u)$ 和 (5.1.1) 的 $J(\boldsymbol{x})$ 有类似结构. 我们亦称 $J(u)$ 为和边值问题 (5.2.1), (5.2.2) 相应的二次泛函或泛函数. 记住这点, 就可针对更一般的椭圆边值问题构造相应的泛函数了.

例如, 对于两点边值问题:

$$Lu=-\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}u}{\mathrm{d}x}\right)+qu=f,\quad x\in(a,b), \tag{5.3.2}$$

$$(5.3.3) \qquad u(a)=0, \quad u'(b)=0,$$

其中 $p\in C^1(\overline{I})$ (一次连续可微函数空间), $p(x)\geqslant \min\limits_{x\in I}p(x)=p_{\min}>0, q\in C(\overline{I})$, $q\geqslant 0, f\in H^0(I), \overline{I}=[a,b]$, 可类似地构造泛函数

$$\begin{aligned} J(u) &= \frac{1}{2}(Lu,u)-(f,u) \\ &= -\frac{1}{2}\int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}u}{\mathrm{d}x}\right)u\mathrm{d}x+\frac{1}{2}\int_a^b qu^2\mathrm{d}x-\int_a^b fu\mathrm{d}x. \end{aligned}$$

对右端第一项施行分部积分, 并用边值条件 (5.3.3) 代入, 得

$$-\int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}u}{\mathrm{d}x}\right)u\mathrm{d}x = -\,p\frac{\mathrm{d}u}{\mathrm{d}x}u\Big|_a^b+\int_a^b p\left(\frac{\mathrm{d}u}{\mathrm{d}x}\right)^2\mathrm{d}x=\int_a^b p\left(\frac{\mathrm{d}u}{\mathrm{d}x}\right)^2\mathrm{d}x.$$

令

$$(5.3.4) \qquad a(u,v)=\int_a^b\left[p\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}v}{\mathrm{d}x}+quv\right]\mathrm{d}x,$$

便得

$$(5.3.5) \qquad J(u)=\frac{1}{2}a(u,u)-(f,u)$$

(参看 (5.1.1)).

设 H_E^1 为 H^1 中满足左边值条件 $u(a)=0$ 的函数组成的子空间. 考虑和 (5.3.2), (5.3.3) 相应的变分问题: 求 $u_*\in H_E^1$ 使

$$(5.3.6) \qquad J(u_*)=\min_{u\in H_E^1}J(u).$$

由 (5.3.4) 定义的 $a(u,v)$ 十分重要, 它在今后的讨论中将起关键作用. 显然 $a(u,v)$ 分别对 u,v 都是线性泛函, 即

$$\begin{aligned} a(c_1u_1+c_2u_2,v)&=c_1a(u_1,v)+c_2a(u_2,v), \\ a(u,c_1v_1+c_2v_2)&=c_1a(u,v_1)+c_2a(u,v_2). \end{aligned}$$

c_1,c_2 是常数, 所以称为双线性泛函或双线性形式. 在讨论极值问题 (5.3.6) 之前, 我们先导出双线性形式 $a(u,v)$ 的几个基本性质.

首先 $a(u,v)$ 是**对称形式**, 即

$$a(u,v)=a(v,u), \text{ 对任意 } u, \quad v\in H^1(I).$$

$a(u,v)$ 的对称性是由微分算子 L 的对称性决定的. 实际上, 设 $u,v\in C^2(I)$, 且满足边值条件 (5.3.3), 则

$$(Lu,v)=\int_a^b\left[-\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}u}{\mathrm{d}x}\right)v+quv\right]\mathrm{d}x \\ =\int_a^b\left[p\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}v}{\mathrm{d}x}+quv\right]\mathrm{d}x. \tag{5.3.7}$$

对调 u,v 后, 等式右端不变, 所以

$$(Lu,v)=(Lv,u)=(u,Lv). \tag{5.3.8}$$

如此的 L 称为**对称算子**.

其次

$$a(u,u)=\int_a^b\left[p\left(\frac{\mathrm{d}u}{\mathrm{d}x}\right)^2+qu^2\right]\mathrm{d}x\geqslant p_{\min}\int_a^b\left(\frac{\mathrm{d}u}{\mathrm{d}x}\right)^2\mathrm{d}x. \tag{5.3.9}$$

如果注意到任一 $u\in H_E^1$ 可表为

$$u(x)=\int_a^x u'(t)\mathrm{d}t,$$

则由 Schwarz 不等式,

$$\int_a^b|u|^2\mathrm{d}x\leqslant\frac{1}{2}(b-a)^2\int_a^b|u'(t)|^2\mathrm{d}t,$$

因而

$$\int_a^b|u'|^2\mathrm{d}x=\frac{1}{2}\left[\int_a^b|u'|^2\mathrm{d}x+\int_a^b|u'|^2\mathrm{d}x\right] \\ \geqslant\frac{1}{(b-a)^2}\int_a^b|u|^2\mathrm{d}x+\frac{1}{2}\int_a^b|u'|^2\mathrm{d}x \\ \geqslant\overline{\gamma}\|u\|_1^2, \tag{5.3.10}$$

其中 $\overline{\gamma}=\min\left(\frac{1}{2},\frac{1}{(b-a)^2}\right)>0$. 形如 (5.3.10) 的不等式称为 **Poincaré** 不等式. 联结 (5.3.9), (5.3.10) 并令 $\gamma=\overline{\gamma}p_{\min}$, 得

$$a(u,u)\geqslant\gamma\|u\|_1^2,\quad 对任意\ u\in H_E^1. \tag{5.3.11}$$

我们称满足不等式 (5.3.11) 的双线性形式为**正定的**. 特别当 $u\in C^2(\overline{I})$ 且满足边值条件 (5.3.3) 时, 由 (5.3.8), (5.3.11) 得

$$(Lu,u)\geqslant\gamma\|u\|_1^2.$$

因此也说 L 是**正定算子**.

最后, 由 Schwarz 不等式知 $a(u,v)$ 满足不等式

$$|a(u,v)| \leqslant M\|u\|_1\|v\|_1, \quad u,v \in H^1(I), \tag{5.3.12}$$

M 是与 u,v 无关的常数. 称 (5.3.12) 为**连续性条件**.

现在回到变分问题 (5.3.6). 像 §1 那样, 任取 $u \in H_E^1$, 考虑实变量 λ 的函数

$$\begin{aligned}\varphi(\lambda) &= J(u_* + \lambda v) \\ &= \frac{1}{2}a(u_* + \lambda v, u_* + \lambda v) - (f, u_* + \lambda v) \\ &= \frac{1}{2}a(u_*, u_*) + \frac{\lambda}{2}[a(u_*, v) + a(v, u_*)] + \frac{\lambda^2}{2}a(v,v) - (f, u_*) - \lambda(f,v).\end{aligned}$$

由 $a(u_*, v)$ 的对称性, 得

$$\varphi(\lambda) = J(u_*) + \lambda[a(u_*, v) - (f,v)] + \frac{\lambda^2}{2}a(v,v). \tag{5.3.13}$$

今证下列变分原理.

定理 3.1　设 $f \in C(I), u_* \in C^2$ 是边值问题 (5.3.2), (5.3.3) 的解, 则 u_* 使 $J(u)$ 达到极小值. 反之, 若 $u_* \in C^2 \bigcap H_E^1$ 使 $J(u)$ 达到极小值, 则 u_* 是边值问题 (5.3.2), (5.3.3) 的解.

证明　注意当 $u_* \in C^2 \bigcap H_E^1, v \in H_E^1$ 时

$$\begin{aligned}a(u_*, v) - (f,v) &= \int_a^b \left[p\frac{\mathrm{d}u_*}{\mathrm{d}x}\frac{\mathrm{d}v}{\mathrm{d}x} + qu_*v - fv\right]\mathrm{d}x \\ &= p\frac{\mathrm{d}u_*}{\mathrm{d}x}v\Big|_a^b + \int_a^b \left[-\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}u_*}{\mathrm{d}x}\right) + qu_* - f\right]v\mathrm{d}x \\ &= \int_a^b [Lu_* - f]v\mathrm{d}x + p(b)u_*'(b)v(b).\end{aligned} \tag{5.3.14}$$

如果 u_* 是边值问题 (5.3.2), (5.3.3) 的解, 则 $Lu_* - f = 0, u_*'(b) = 0$, 从而

$$\varphi'(0) = a(u_*, v) - (f,v) = 0, \quad \text{对任意 } v \in H_E^1,$$

(注意 (5.3.14)). 由 (5.3.13) 及 $a(u,v)$ 的正定性, 当 $\lambda \neq 0, v \neq 0$ 时

$$J(u_* + \lambda v) = J(u_*) + \frac{\lambda^2}{2}a(v,v) > J(u_*).$$

这说明 u_* 使 $J(u)$ 达到极小.

反之, 若 u_* 使 $J(u)$ 达到极小值, 则由 (5.3.13), (5.3.14) 得

$$\begin{aligned}\varphi'(0) &= a(u_*, v) - (f, v) \\ &= \int_a^b [Lu_* - f]v\mathrm{d}x + p(b)u_*'(b)v(b) = 0 \quad \text{对任意 } v \in H_E^1.\end{aligned} \tag{5.3.15}$$

特别取 $v \in C_0^\infty(I)$, 则

$$\int_a^b [Lu_* - f]v\mathrm{d}x = 0, \quad \text{对任意 } v \in C_0^\infty(I).$$

根据变分法基本引理,u_* 满足方程

$$Lu_* - f = 0.$$

于是 (5.3.15) 化为

$$p(b)u'_*(b)v(b) = 0, \quad \text{对任意 } v \in H_E^1.$$

注意 $p(b) > 0$, 取 $v(x) = x - a$, 则 $v \in H_E^1$, 且 $v(b) > 0$, 可见 u_* 必须满足右边值条件

$$u_*'(b) = 0.$$

因为在力学、物理中, 二次泛函 $J(u)$ 表示能量, 所以也称定理 3.1 为**极小位能原理**. 应当注意的是, 我们仅就二次连续可微解 u_*(称为古典解) 建立了边值问题和变分问题的等价性. 对于非光滑函数 u_*, 说它是边值问题的解就没有意义了. 但是许多物理、力学现象, 必须用非光滑函数才能描述它. 比如前面所举弦平衡的例子, 若作用于弦的外力是集中荷载, 则弦的平衡曲线 $u = u(x)$ 不再有连续的二阶导数, 某些点甚至是没有导数的 “尖点”, 这时 $u(x)$ 在古典意义下不可能是 (5.3.2), (5.3.3) 的解. 而能量 $J(u)$ 的表达式是积分式 (5.3.9), 被积函数只含 u 的一阶导数, 只要 u 连续且按段连续可微, 则 $J(u)$ 有意义. 因此变分问题 (5.3.6) 允许非光滑解 $u_* = u_*(x)$, 称之为两点边值问题 (5.3.2), (5.3.3) 的**广义解**.

边值问题可能有广义解但没有古典解. 定理 3.1 告诉我们, 当边值问题存在古典解时, 它一定是广义解. 反之, 若广义解存在且二次连续可微, 则广义解就是古典解.

按照变分法, 我们称 (5.3.2) 是和泛函数 $J(u)$ 相关的 **Euler 方程**.

注 3.1 同理可证边值问题 (5.2.1), (5.2.2) 和变分问题 (5.2.4)′ 等价.

从定理 3.1 知道, 左边值条件 $u(a) = 0$ 和右边值条件 $u'(b) = 0$ 有重要区别. 前者必须强加在变分问题所在的函数类上, 称为**强制边值条件**或**本质边值条件**.

后者不必对函数类作为条件提出, 只要函数 $u_*(x)$ 使 $J(u)$ 取极小值, 则它必然满足该条件, 因此称为**自然边值条件**. 在数值求解边值问题时, 区别这两类条件很重要, 这是从变分问题出发构造数值方法的一个优点.

3.2 虚功原理

以 v 乘 (5.3.2) 两端, 沿区间 $[a,b]$ 积分, 得

$$\int_a^b (Lu-f)v\mathrm{d}x = \int_a^b \left[-\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}u}{\mathrm{d}x}\right)v + quv - fv\right]\mathrm{d}x = 0. \tag{5.3.16}$$

利用分部积分和关于 u,v 的边值条件 (5.3.3), 则

$$-\int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}u}{\mathrm{d}x}\right)v\mathrm{d}x = -\left.p\frac{\mathrm{d}u}{\mathrm{d}x}v\right|_a^b + \int_a^b p\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}v}{\mathrm{d}x}\mathrm{d}x = \int_a^b p\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}v}{\mathrm{d}x}\mathrm{d}x.$$

以此代入到 (5.3.16) 式, 得

$$\int_a^b \left[p\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}v}{\mathrm{d}x} + quv - fv\right]\mathrm{d}x = 0.$$

若注意到双线性形式 $a(u,v)$ 的表达式 (5.3.4), 则上式可写成

$$a(u,v) - (f,v) = 0. \tag{5.3.17}$$

这也是边值问题 (5.3.2), (5.3.3) 的变分形式, 其确切提法将在定理 3.2 中给出.

对 $u\in C^2\bigcap H_E^1, v\in H_E^1$, 根据 (5.3.14), 方程 (5.3.17) 左端

$$a(u,v)-(f,v) = \int_a^b [Lu-f]v\mathrm{d}x + p(b)u'(b)v(b).$$

假若 u 是边值问题 (5.3.2), (5.3.3) 的解, 则对任意 $v\in H_E^1, u$ 满足 (5.3.17). 反之, 若对任意 $v\in H_E^1, u\in H_E^1$ 满足 (5.3.17), 则可按定理 3.1 的证法, 推出 u 是边值问题 (5.3.2), (5.3.3) 的解, 于是有

定理 3.2 设 $u\in C^2$, 则 u 是边值问题 (5.3.2),(5.3.3) 的解的充要条件是:$u\in H_E^1$ 且满足变分方程

$$a(u,v)-(f,v)=0, \text{对任意 } v\in H_E^1. \tag{5.3.18}$$

在力学中, (5.3.18) 左端表示虚功, 所以也称定理 3.2 为**虚功原理**. 当 u 是边值问题的古典解时, 它也是变分方程 (5.3.18) 的解. 像位能原理一样, 变分方程 (5.3.18) 还允许非古典解, 我们称这样的解为边值问题的**广义解**.

虚功原理比位能原理更具有一般性, 它不仅适用于**对称正定算子方程** (相当于力学中的**保守场方程**), 而且也适用于**非对称正定算子方程** (**非保守场方程**). 实际上, 定理 3.2 可直接推广为

定理 3.3 设 $u \in C^2$, 则 u 满足

$$\begin{cases} Lu = -\dfrac{\mathrm{d}}{\mathrm{d}x}\left(p\dfrac{\mathrm{d}u}{\mathrm{d}x}\right) + r\dfrac{\mathrm{d}u}{\mathrm{d}x} + qu = f, \\ u(a) = 0, \quad u'(b) = 0 \end{cases}$$

的充要条件是: $u \in C^2 \bigcap H_E^1$ 且满足变分方程:

$$\begin{cases} a(u,v) - (f,v) = 0, \quad \text{任意 } v \in H_E^1, \\ a(u,v) = \displaystyle\int_a^b \left[p\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}v}{\mathrm{d}x} + r\frac{\mathrm{d}u}{\mathrm{d}x}v + quv\right]\mathrm{d}x, \end{cases}$$

其中 $p \in C^1, p_{\min} > 0, r, q \in C, f \in L^2$.

此时双线性形式 $a(u,v)$ 非对称正定, 除非 $r \equiv 0, q \geqslant 0$.

习　　题

1. 证明非齐次两点边值问题

$$\begin{cases} Lu = -\dfrac{\mathrm{d}}{\mathrm{d}x}\left(p\dfrac{\mathrm{d}u}{\mathrm{d}x}\right) + qu = f, \quad a < x < b, \\ u(a) = \alpha, \quad u'(b) = \beta, \end{cases} \tag{5.3.19}$$

与下列变分问题等价: 求 $u_* \in H^1, u_*(a) = \alpha$, 使

$$J(u_*) = \min_{\substack{u \in H^1 \\ u(a)=\alpha}} J(u),$$

其中

$$J(u) = \frac{1}{2}a(u,u) - (f,u) - p(b)\beta u(b),$$

而 $a(u,v)$ 如 (5.3.4) (提示: 先把边值条件齐次化).

2. 就边值问题 (5.3.19) 建立虚功原理.

3. 试建立与边值问题

$$\begin{cases} Lu = \dfrac{\mathrm{d}^4 u}{\mathrm{d}x^4} + u = f, \quad a < x < b, \\ u(a) = u'(a) = 0, \quad u(b) = u'(b) = 0 \end{cases}$$

等价的变分问题.

§4 二阶椭圆边值问题

4.1 极小位能原理

作为模型, 考虑 Poisson 方程的第一边值问题:

$$-\Delta u = f(x,y), (x,y) \in G, \tag{5.4.1}$$

$$u|_\Gamma = 0, \tag{5.4.2}$$

边界 Γ 为分段光滑曲线. 仿照 §1, §3, 作泛函数

$$\begin{aligned} J(u) &= \frac{1}{2}(-\Delta u, u) - (f, u) \\ &= \frac{1}{2}\iint\limits_G (-\Delta u)u\mathrm{d}x\mathrm{d}y - \iint\limits_G fu\mathrm{d}x\mathrm{d}y. \end{aligned} \tag{5.4.3}$$

利用 Green 公式, 我们得

$$\iint\limits_G (-\Delta u)v\mathrm{d}x\mathrm{d}y = \iint\limits_G \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)\mathrm{d}x\mathrm{d}y - \int_\Gamma \frac{\partial u}{\partial n}v\mathrm{d}s, \tag{5.4.4}$$

其中 $\boldsymbol{n}$ 表示曲线边界 Γ 的单位外法向, $\frac{\partial u}{\partial n}$ 是 u 沿 $\boldsymbol{n}$ 的方向导数. 公式 (5.4.4) 很重要, 称为 **Green 第一公式**. 若 u, v 满足边值条件 (5.4.2), 则

$$\iint\limits_G (-\Delta u)v\mathrm{d}x\mathrm{d}y = \iint\limits_G \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)\mathrm{d}x\mathrm{d}y. \tag{5.4.5}$$

定义双线性形式

$$a(u, v) = \iint\limits_G \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)\mathrm{d}x\mathrm{d}y. \tag{5.4.6}$$

由 (5.4.3), (5.4.5) 和 (5.4.6), 可将泛函数 $J(u)$ 写成 (参看 (5.1.1) 和 (5.3.5)):

$$J(u) = \frac{1}{2}a(u, u) - (f, u). \tag{5.4.7}$$

在力学上, $J(u)$ 表示位能.

从 (5.4.6), (5.4.7) 知道, 只要 $u \in H^1(G), f \in L^2(G)$, 则 $J(u)$ 有意义. 此外还要求 u 满足第一边值条件 (5.4.2). 以下用 $H_0^1(G)$ 表示 $H^1(G)$ 中一切满足 (5.4.2) 的函数组成的子空间.

现在提如下变分问题: 求 $u_* \in H_0^1(G)$, 使

$$J(u_*) = \min_{u \in H_0^1(G)} J(u). \tag{5.4.8}$$

为了建立边值问题 (5.4.1), (5.4.2) 和变分问题 (5.4.8) 的等价性, 先讨论双线性形式 $a(u, v)$ 的两个基本性质.

(1) **对称性**. 显然

$$a(u, v) = a(v, u), \quad 对任意\ u, \quad v \in H^1(G).$$

(2) **正定性**. 对于 $u \in H_0^1(G)$,

$$a(u,u) = \iint\limits_G \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right] \mathrm{d}x\mathrm{d}y = ||u_x||^2 + ||u_y||^2.$$

与 (5.3.10) 类似, 可建立 Poincare 不等式:

$$(5.4.9) \qquad ||u_x||^2 + ||u_y||^2 \geqslant \gamma ||u||_1^2, \quad u \in H_0^1(G),$$

其中 $\gamma > 0$ 是和 u 无关的常数. 于是

$$(5.4.10) \qquad a(u,u) \geqslant \gamma ||u||_1^2, \quad u \in H_0^1(G).$$

这说明 $a(u,v)$ 正定.

其次, 对于 $u, v \in H^1(G), a(u,v)$ 满足不等式 (连续性条件)

$$(5.4.11) \qquad |a(u,v)| \leqslant ||u||_1 ||v||_1.$$

由于 $a(u,v)$ 对称正定, 也称 $-\Delta$ 为对称正定算子 (参看 (5.4.5), (5.4.6)).

对于 $u_*, u \in H_0^1(G)$, 考虑实参数 λ 的函数

$$\varphi(\lambda) = J(u_* + \lambda u).$$

利用 $a(u,v)$ 的对称性, 可知

$$(5.4.12) \qquad \varphi(\lambda) = J(u_*) + \lambda[a(u_*, u) - (f,u)] + \frac{\lambda^2}{2} a(u,u).$$

它和 (5.3.13) 有完全相同的形式.

若进一步假定 $u_* \in C^2(\overline{G}) \bigcap H_0^1(G)$, 则由 (5.4.5), (5.4.6) 得出

$$(5.4.13) \qquad a(u_*, u) - (f,u) = (-\Delta u_* - f, u).$$

设 u_* 是边值问题 (5.4.1), (5.4.2) 的解, 则

$$\varphi'(0) = a(u_*, u) - (f,u) = (-\Delta u_* - f, u) = 0, \quad 对任意\ u \in H_0^1(G),$$

从而

$$J(u_* + \lambda u) = J(u_*) + \frac{\lambda^2}{2} a(u,u) > J(u_*),$$

对任意 $u \neq 0, u \in H_0^1(G), \lambda \neq 0$. 这说明 u_* 使 $J(u)$ 达到极小.

与定理 3.1 类似, 可证明使 $J(u)$ 达到极小的 u_*, 当其属于 $C^2(\overline{G}) \bigcap H_0^1(G)$ 时, 必为 (5.4.1), (5.4.2) 的解.

于是得

定理 4.1　设 $u_* \in C^2(\overline{G})$ 是边值问题 (5.4.1), (5.4.2) 的解, 则 u_* 使 $J(u)$ 达到极小. 反之, 若 $u_* \in C^2(\overline{G}) \bigcap H_0^1(G)$ 使 $J(u)$ 达到极小, 则 u_* 是边值问题 (5.4.1), (5.4.2) 的解 (古典解).

由于 $J(u)$ 在力学、物理学中表示能量, 所以也称定理 4.1 为**极小位能原理**. 注意定理 4.1 要求 $u_* \in C^2(\overline{G})$, 而变分问题 (5.4.8) 还允许不属于 $C^2(\overline{G})$ 的解, 称之为**边值问题的广义解**.

注 4.1　若代替 (5.4.2) 的是非齐边值条件

$$u|_\Gamma = \varphi(x,y), \quad \varphi \in C^1(\Gamma), \tag{5.4.14}$$

则取一特定函数 $u_0 \in C^2(\overline{G}), u_0|_\Gamma = \varphi$, 令 $v = u - u_0$, 则 v 满足方程:

$$-\Delta v = f + \Delta u_0$$

和齐次边值条件 (5.4.2). 构造 v 的二次泛函

$$\begin{aligned}
\widehat{J}(v) &= \frac{1}{2}\iint\limits_G \left[\left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2\right] \mathrm{d}x\mathrm{d}y - \iint\limits_G (f + \Delta u_0) v \mathrm{d}x\mathrm{d}y \\
&= \frac{1}{2}\iint\limits_G \left[\left(\frac{\partial u}{\partial x} - \frac{\partial u_0}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y} - \frac{\partial u_0}{\partial y}\right)^2\right] \mathrm{d}x\mathrm{d}y - \\
&\quad \iint\limits_G (f + \Delta u_0)(u - u_0) v \mathrm{d}x\mathrm{d}y \\
&= \frac{1}{2}\iint\limits_G \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 - 2fu\right] \mathrm{d}x\mathrm{d}y - \iint\limits_G \left[\frac{\partial u}{\partial x}\frac{\partial u_0}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial u_0}{\partial y}\right] \mathrm{d}x\mathrm{d}y - \\
&\quad \iint\limits_G \Delta u_0 u \mathrm{d}x\mathrm{d}y + \text{常数}.
\end{aligned}$$

由 (5.4.4)

$$\begin{aligned}
&-\iint\limits_G \Delta u_0 u \mathrm{d}x\mathrm{d}y - \iint\limits_G \left[\frac{\partial u}{\partial x}\frac{\partial u_0}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial u_0}{\partial y}\right] \mathrm{d}x\mathrm{d}y \\
= &-\int\limits_\Gamma u \frac{\partial u_0}{\partial \boldsymbol{n}} \mathrm{d}s = -\int\limits_\Gamma \varphi \frac{\partial u_0}{\partial \boldsymbol{n}} \mathrm{d}s = \text{常数},
\end{aligned}$$

足见

$$\widehat{J}(v) = J(u) + \text{常数}.$$

由此可见, 变分问题

$$\widehat{J}(v_*) = \min_{v \in H_0^1} \widehat{J}(v)$$

和

$$J(u_*) = \min_{\substack{u \in H^1(G) \\ u|_\Gamma = \varphi}} J(u) \tag{5.4.15}$$

等价, 且 $v_* = u_* - u_0$.

据定理 4.1, 非齐次边值问题 (5.4.1), (5.4.14) 与变分问题 (5.4.15) 等价.

4.2 自然边值条件

考虑第二、第三边值条件

$$\left.\frac{\partial u}{\partial \boldsymbol{n}} + \alpha u\right|_\Gamma = 0, \quad \alpha \geqslant 0. \tag{5.4.16}$$

利用公式 (5.4.4),

$$\begin{aligned}
J(u) &= \frac{1}{2}(-\Delta u, u) - (f, u) \\
&= \frac{1}{2}\iint\limits_G \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right] \mathrm{d}x\mathrm{d}y - \frac{1}{2}\int\limits_\Gamma \frac{\partial u}{\partial \boldsymbol{n}} u \mathrm{d}s - \iint\limits_G f u \mathrm{d}x\mathrm{d}y \\
&= \frac{1}{2}\iint\limits_G \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right] \mathrm{d}x\mathrm{d}y + \frac{1}{2}\int\limits_\Gamma \alpha u^2 \mathrm{d}s - \iint\limits_G f u \mathrm{d}x\mathrm{d}y.
\end{aligned}$$

令

$$a(u, v) = \iint\limits_G \left[\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right] \mathrm{d}x\mathrm{d}y + \int\limits_\Gamma \alpha u v \mathrm{d}s, \tag{5.4.17}$$

则

$$J(u) = \frac{1}{2}a(u, u) - (f, u). \tag{5.4.18}$$

设 $u_* \in C^2(\overline{G}), u \in H^1(G)$. 由公式 (5.4.4), 我们有

$$a(u_*, u) - (f, u) = (-\Delta u_* - f, u) + \int\limits_\Gamma \left(\alpha u_* + \frac{\partial u_*}{\partial \boldsymbol{n}}\right) u \mathrm{d}s. \tag{5.4.19}$$

今考虑实变量 λ 的函数

$$\varphi(\lambda) = J(u_* + \lambda u), \quad u_*, u \in H^1.$$

直接计算, 可得形如 (5.4.12) 的展式:

$$\varphi(\lambda) = J(u_*) + \lambda[a(u_*, u) - (f, u)] + \frac{\lambda^2}{2}a(u, u), \tag{5.4.20}$$

其中 $a(u_*, u)$ 由 (5.4.17) 定义. 与定理 3.1 的证法类似, 可以证明

定理 4.2　边值问题 (5.4.1), (5.4.16) 的解 $u_* \in C^2(\overline{G})$ 是下列变分问题的解: 求 $u_* \in H_0^1(G)$, 使

$$J(u_*) = \min_{u \in H^1} J(u). \tag{5.4.21}$$

反之, 变分问题 (5.4.21) 的解 u_* 若属于 $C^2(\overline{G})$, 则也是边值问题 (5.4.1), (5.4.16) 的解.

若 $u_* \in H^1(\overline{G})$ 是 (5.4.21) 的解, 则称之为**边值问题的广义解**.

值得指出的是, 变分问题 (5.4.21) 并不要求 u 满足任何边值条件, 而它的解 u_* 却自动满足 (5.4.16), 这是第二、三边值条件与第一边值条件的一个重大差别. 像两点边值问题一样, 我们称第一边值条件为**本质边值条件**, 第二、三边值条件为**自然边值条件**.

4.3　虚功原理

像 §3 那样, 同样可以建立第一、第二、第三边值问题的虚功原理. 为叙述统一, 我们考虑 Poisson 方程 (5.4.1) 的混合边值问题. 如图 4.1, 设边界 Γ 分成互不相交的两部分: Γ_1 和 Γ_2. 在 Γ_1 上满足第一边值条件:

$$u|_{\Gamma_1} = 0, \tag{5.4.22}$$

在 Γ_2 上满足第二或第三边值条件:

$$\left.\frac{\partial u}{\partial \boldsymbol{n}} + \alpha u\right|_{\Gamma_2} = 0, \quad \alpha \geqslant 0. \tag{5.4.23}$$

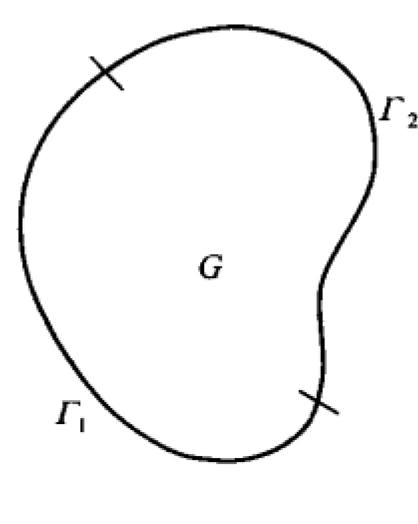


图 4.1

以 v 乘 (5.4.1) 两端并在 G 上积分, 得

$$\int_G [(-\Delta u)v - fv]\mathrm{d}x\mathrm{d}y = 0. \tag{5.4.24}$$

利用公式 (5.4.4) 及关于 u,v 的边值条件 (5.4.22), (5.4.23) 得

$$\tag{5.4.25} \begin{aligned}\iint\limits_G (-\Delta u)v\mathrm{d}x\mathrm{d}y &= \iint\limits_G \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)\mathrm{d}x\mathrm{d}y - \int\limits_\Gamma \frac{\partial u}{\partial \boldsymbol{n}}v\mathrm{d}s\\ &= \iint\limits_G \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)\mathrm{d}x\mathrm{d}y + \int\limits_{\Gamma_2} \alpha uv\mathrm{d}s.\end{aligned}$$

定义双线性形式

$$\tag{5.4.26} a(u,v) = \iint\limits_G \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)\mathrm{d}x\mathrm{d}y + \int\limits_{\Gamma_2} auv\mathrm{d}s,$$

则可将 (5.4.24) 写成

$$\tag{5.4.27} a(u,v)-(f,v)=0.$$

以 $H_E^1(G)$ 表示 $H^1(G)$ 中满足第一边值条件 (5.4.22) 的函数组成的子空间. 今提如下变分问题: 求 $u\in H_E^1(G)$, 使 u 对一切 $v\in H_E^1(G)$ 满足 (5.4.27).

设 $u\in C^2(\overline{G}), v\in H_E^1(G)$, 则由 (5.4.4) 得

$$\tag{5.4.28} a(u,v)-(f,v) = \iint\limits_G [-\Delta u - f]v\mathrm{d}x\mathrm{d}y + \int\limits_{\Gamma_2}\left(\frac{\partial u}{\partial \boldsymbol{n}}+\alpha u\right)v\mathrm{d}s.$$

与定理 3.1 的证法类似, 可推出

定理 4.3 设 $u\in C^2(\overline{G})$, 则 u 满足 (5.4.1), (5.4.22), (5.4.23) 的充要条件是: $u\in H_E^1$ 且对任意 $v\in H_E^1$ 满足变分方程 (5.4.27).

因为 (5.4.27) 左端在力学中表示虚功, 故亦称定理 4.3 为**虚功原理**. 和边值问题不同, 变分方程 (5.4.27) 允许有不属于 $C^2(G)$ 的解, 称为**边值问题的广义解**.

从定理 4.3 看出, 边值条件 (5.4.22) 和 (5.4.23) 有重要差别, 前者为本质边值条件, 后者为自然边值条件.

正如 §3 指出过的, 虚功原理较极小位能原理应用更广, 它不必要求边值问题对称正定.

习　题

1. 设 $u\in C(\overline{G})$ 满足

$$\iint\limits_G u\varphi\mathrm{d}x\mathrm{d}y = 0,\quad 对任意\ \varphi\in C_0^\infty(G),$$

试证 $u\equiv 0$.

2. 证明定理 4.1 的第二部分.

3. 试就 Poisson 方程 (5.4.1) 的非齐次边值条件

$$\left.\frac{\partial u}{\partial \boldsymbol{n}}+\alpha u\right|_{\Gamma}=\beta, \quad \alpha \geqslant 0, \tag{5.4.29}$$

导出等价的变分问题.

4. 试就椭圆方程第一边值问题:

$$-\nabla(k\nabla u)+\sigma u=f, \quad (x,y)\in G, \quad u|_{\Gamma}=g \tag{5.4.30}$$

建立等价的极小位能原理和虚功原理, 其中 $k=k(x,y)\in C^1(\overline{G}), \min\limits_{\overline{G}} k>0, \sigma\in C(\overline{G}), \sigma\geqslant 0, f\in L^2(G), g\in C(\Gamma)$, 而

$$\nabla(k\nabla u)=\frac{\partial}{\partial x}\left(k\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial y}\left(k\frac{\partial u}{\partial y}\right).$$

§5 Ritz–Galerkin 方法

前面各节讨论了如何化边值问题为等价的变分问题, 本节讨论如何解相应的变分问题. 必须指出, 除少数特殊情形外, 一般不可能求得问题的准确解, 因此需要各种近似或数值解法. Ritz–Galerkin 方法是最重要的一种解法, 它是以后要讨论的有限元法的基础.

用 U 表示 H_0^1, H_E^1, H^1 等 Sobolev 空间, $H=H^0$ 是 L^2 空间. L 代表 §3, §4 中的微分算子 (二阶常微分或偏微分算子). $a(u,v)$ 是由 L 及边值条件决定的双线性形式, 它由 (Lu,v) 经过分部积分并代入边值条件后得到. 得出 $a(u,v)$ 的表达式后, u,v 就无需满足自然边值条件了, 但本质边值条件仍需满足, 就是说, u,v 应属于空间 U. 前已证明, $a(u,v)$ 是对称正定双线性形式, 即满足

$$a(u,v)=a(v,u), \quad \text{对任意 } u,v\in U, \tag{5.5.1}$$

$$a(u,u)\geqslant \gamma\|u\|_1^2, \quad \text{对任意 } u\in U, \tag{5.5.2}$$

其中 $\gamma>0$ 是与 u 无关的常数.

此外, $a(u,v)$ 还满足连续性条件

$$|a(u,v)|\leqslant M\|u\|_1\|v\|_1, \quad u,\ v\in H^1 \tag{5.5.3}$$

(参看 (5.3.12), (5.4.11)).

设 $f \in H$, 则 §1, §3 和 §4 的二次泛函可统一写成形式:

$$J(u) = \frac{1}{2}a(u,u) - (f,u).$$

于是边值问题 $Lu = f$ 等价于求 $u \in U$, 使

$$J(u) = \min_{v \in U} J(v). \tag{5.5.4}$$

这就是极小位能原理.

边值问题的另一变分形式是: 求 $u \in U$, 使

$$a(u,v) = (f,v), \text{对任意 } v \in U. \tag{5.5.5}$$

这就是虚功原理. 虚功原理并不要求 $a(u,v)$ 对称正定.

变分问题 (5.5.4) 和通常的极值问题 (§1) 比较, 主要困难是在无穷维空间 U 上求泛函的极小值. Ritz Galerkin 方法的基本思想在于用有穷维空间近似代替无穷维空间, 从而化成多元二次函数的极值问题 (参看 §1). 关键是如何选取有穷维空间.

设 U_n 是 U 的 n 维子空间, $\varphi_1, \varphi_2, \cdots, \varphi_n$ 是 U_n 的一组基底, 称为**基函数**. 则 U_n 中任一函数 u_n 可表为

$$u_n = \sum_{i=1}^{n} c_i \varphi_i. \tag{5.5.6}$$

Ritz 法的目标是: 选取系数 c_i, 使 $J(u_n)$ 取极小值. 注意

$$J(u_n) = \frac{1}{2}a(u_n,u_n) - (f,u_n) = \frac{1}{2}\sum_{i,j=1}^{n} a(\varphi_i,\varphi_j)c_i c_j - \sum_{j=1}^{n} c_j (f,\varphi_j)$$

是 $c_1, c_2, \cdots, c_n$ 的二次函数, $a(\varphi_i,\varphi_j) = a(\varphi_j,\varphi_i)$. 如 §1, 令

$$\frac{\partial J(u_n)}{\partial c_j} = 0,$$

即知 $c_1, c_2, \cdots, c_n$ 满足

$$\sum_{i=1}^{n} a(\varphi_i,\varphi_j)c_i = (f,\varphi_j), \quad j = 1,2,\cdots,n. \tag{5.5.7}$$

这是线代数方程组, 求出 c_i 后, 代到 (5.5.6), 就得出近似解 u_n, 这就是熟知的 Ritz 法. 用 Ritz 法求得的 u_n 在空间 U_n 是最佳的, 就是说, 在 U_n 中的所有元素中, u_n 使 $J(u)$ 达到极小值.

Galerkin 法也是求形如 (5.5.6) 的近似解, 但要求系数 c_i 使 u_n 关于 $v \in U_n$ 满足 (5.5.5), 即

$$a(u_n, v) = (f, v), \quad 对任意\ v \in U_n, \tag{5.5.8}$$

或 (取 $v = \phi_j, 1 \leqslant j \leqslant n$)

$$\sum_{i=1}^{n} a(\varphi_i, \varphi_j) c_i = (f, \varphi_j), \quad j = 1, 2, \cdots, n.$$

这和 Ritz 法导出的方程组 (5.5.7) 相同, 因此习惯上称 (5.5.7) 为 **Ritz–Galerkin 方程**.

尽管 Ritz 法和 Galerkin 法导出的近似解 u_n 及计算方法完全一样, 但二者的基础不同. Ritz 法基于极小位能原理, 而 Galerkin 法基于虚功原理, 所以 Galerkin 法较 Ritz 法应用更广, 方法推导也更直接. 仅当 $a(u, v)$ 对称正定时两者才一致; 否则, 只能用 Galerkin 法, 而不能用 Ritz 法.Ritz 法的优点是: 力学意义更明显 (尤其是特征值问题), 理论基础比较容易建立.

现在考察 Ritz–Galerkin 方程 (5.5.7) 的系数矩阵

$$\boldsymbol{A} = \begin{bmatrix} a(\varphi_1, \varphi_1) & a(\varphi_2, \varphi_1) & \cdots & a(\varphi_n, \varphi_1) \\ \vdots & \vdots & & \vdots \\ a(\varphi_1, \varphi_n) & a(\varphi_2, \varphi_n) & \cdots & a(\varphi_n, \varphi_n) \end{bmatrix}. \tag{5.5.9}$$

据 (5.5.1), $a(\varphi_i, \varphi_j) = a(\varphi_j, \varphi_i)$, 所以 $\boldsymbol{A}$ 对称. 又由 (5.5.2), 对任一非零向量 $(c_1, \cdots, c_n)^{\mathrm{T}}$, 二次型

$$\sum_{i,j=1}^{n} a(\varphi_i, \varphi_j) c_i c_j = a\left(\sum_{i=1}^{n} c_i \varphi_i, \sum_{j=1}^{n} c_j \varphi_j\right) = a(u_n, u_n) > 0,$$

这说明 $\boldsymbol{A}$ 正定. 由此即知 (5.5.7) 唯一可解.

定理 5.1 设 u 是变分问题 (5.5.4) 或 (5.5.5) 的解,u_n 是 Ritz–Galerkin 方程 (5.5.7) 的解, 则有与 u, n 无关的常数 $\beta > 0$, 使误差 $u - u_n$ 满足不等式:

$$\|u - u_n\|_1 \leqslant \beta \inf_{v \in U_n} \|u - v\|_1. \tag{5.5.10}$$

如果 $\{\varphi_i\}_1^\infty$ 于 U 完全, 即 $\{\varphi_i\}_1^\infty$ 的一切可能的线性组合于 U 稠密, 则由 (5.5.10) 可推出收敛性:

$$\lim_{n \to \infty} \|u - u_n\|_1 = 0. \tag{5.5.11}$$

估计式 (5.5.10) 左边的 $||u-u_n||_1$ 表示 u 与 u_n 的距离, 右边的 $\inf\limits_{v\in U_n}||u-v||_1$ 表示 u 和子空间 U_n 的最小距离. 注意 $u_n\in U_n$, 可见估计式 (5.5.10) 是最佳的 (如果不考虑常系数 β).

证明 从 (5.5.10) 推 (5.5.11) 是直接的. 实际上, 对任意 $\varepsilon>0$, 有线性组合 $\sum\limits_{i=1}^{m}d_i\varphi_i$ 使

$$\left\|u-\sum_{i=1}^{m}d_i\varphi_i\right\|_1<\frac{\varepsilon}{\beta}.$$

由 (5.5.10), 当 $n>m$ 时,

$$||u-u_n||_1\leqslant\beta\left\|u-\sum_{i=1}^{m}d_i\varphi_i\right\|_1<\beta\frac{\varepsilon}{\beta}=\varepsilon.$$

这证明了 (5.5.11).

今证 (5.5.10). 因 u, u_n 依次满足方程:

$$a(u,v)=(f,v),\quad \text{对任意 } v\in U,$$
$$a(u_n,\varphi_j)=(f,\varphi_j),\quad j=1,2,\cdots,n,$$

故对任意 $v_n\in U_n$,

$$a(u,v_n)=(f,v_n),\quad a(u_n,v_n)=(f,v_n).$$

两式相减, 得

(5.5.12) $$a(u-u_n,v_n)=0,\quad \text{对任意 } v_n\in U_n.$$

利用 (5.5.2) 和 (5.5.12), 有

(5.5.13) $$\begin{aligned}||u-u_n||_1^2&\leqslant\frac{1}{\gamma}a(u-u_n,u-u_n)=\frac{1}{\gamma}a(u-u_n,u)\\&=\frac{1}{\gamma}a(u-u_n,u-v_n).\end{aligned}$$

由 (5.5.3),

$$a(u-u_n,u-v_n)\leqslant M||u-u_n||_1||u-v_n||_1,$$

因此

$$||u-u_n||_1^2\leqslant\frac{M}{\gamma}||u-u_n||_1||u-v_n||_1.$$

消去共同因子 $||u-u_n||_1$, 并令 $\beta = M/\gamma$, 则得

$$||u-u_n||_1 \leqslant \beta||u-v_n||_1, \quad \text{对任意 } v_n \in U_n.$$

两端关于 $v_n \in U_n$ 取下确界, 即得 (5.5.10).

注 5.1 当 Ritz–Galerkin 法用于非齐边值问题时, 要根据边值条件的两种不同类型 (本质的和自然的) 作相应处理. 对非齐次自然边值条件, 只要适当修改右端即可, 不必对基函数加任何限制. 对于非齐次本质边值条件, 应对它齐次化后再用 Ritz–Galerkin 方法. 例如非齐次边值问题 (5.3.19), 其右端点为自然边值条件, 因此右端应改为 $(f,\varphi_j)+p(b)\beta\varphi_j(b)$ (参看 §3 习题 2,3). 而左端点为本质边值条件, 经齐次化后, u_n 形如

$$u_n(x) = u_0(x) + \sum_{i=1}^{n} c_i\varphi_i(x),$$

其中 $\varphi_i(a)=0$ $(i=1,2,\cdots,n), u_0(x)$ 是满足 $u_0(a)=\alpha$ 的任一已知函数. 相应的 Ritz–Galerkin 方程变成

$$\sum_{i=1}^{n} a(\varphi_i,\varphi_j)c_i = (f,\varphi_j)+p(b)\beta\varphi_j(b) - a(u_0,\varphi_j), \quad j=1,2,\cdots,n. \tag{5.5.14}$$

实际计算时取 $u_0(x)=\alpha\varphi_0(x), \varphi_0(x)$ 是满足 $\varphi_0(a)=1$ 的任一函数 (为使右端点条件保持不变, 要求 $\varphi_0(x)$ 在 b 附近等于 0). 对于二维边值问题, 精确给出 $u_0(x)$ 是困难的, 一般只能用插值法得到 $u_0(x)$ 的近似式.

注 5.2 我们曾经指出, Ritz 法只能用于解对称正定微分算子方程, 而 Galerkin 法则可解更一般的微分方程. 例如两点边值问题 (参看定理 3.3):

$$\begin{cases} Lu = -\dfrac{\mathrm{d}}{\mathrm{d}x}\left(p\dfrac{\mathrm{d}u}{\mathrm{d}x}\right) + r\dfrac{\mathrm{d}u}{\mathrm{d}x} + qu = f, \quad a<x<b, \\ u(a)=0,\ u'(b)=0. \end{cases} \tag{5.5.15}$$

其中 $p\in C^1(\overline{I}), p(x)\geqslant p_{\min}>0, r,q\in C(\overline{I}), f\in L^2(I), I=(a,b)$. 与之相应的双线性形式为

$$a(u,v)=\int_a^b \left[p\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}v}{\mathrm{d}x} + r\frac{\mathrm{d}u}{\mathrm{d}x}v + quv\right]\mathrm{d}x.$$

显然 $a(u,v)$ 非对称正定, 除非 $r\equiv 0, q\geqslant 0$. 因此不能用 Ritz 法解 (5.4.15). 但 Galerkin 法仍然可用, 且导出的线性方程和 (5.5.7) 相同.

至于 Galerkin 方法收敛性和误差估计的研究, 已超出本书范围, 此处从略 (参看 [10]).

例 5.1　用 Ritz–Galerkin 法解边值问题

$$
(5.5.16)\qquad \begin{cases} u''+u=-x, \quad 0<x<1, \\ u(0)=u(1)=0. \end{cases}
$$

此时, $U=H_0^1(I)$ $(I=(0,1)), H=L^2(I)$. 于 $H_0^1(I)$ 取一族基函数 $\varphi_i(x)$ $(i=1,2\cdots)$, 使每一 $\varphi_i(x)$ 满足齐次边值条件, 彼此线性独立, 且构成 H_0^1 的完全系统. 以 $\varphi_1,\varphi_2,\cdots,\varphi_n$ 为基底张开的子空间就是 n 维空间 U_n.

通常有两种选取 φ_i 的方法. 一种是选 φ_i 为三角多项式

$$
\varphi_i(x)=\sin(i\pi x), \quad i=1,2,\cdots,
$$

另一种是取 φ_i 为代数多项式

$$
\varphi_i(x)=\omega(x)x^{i-1}, \quad i=1,2,\cdots.
$$

为使 φ_i 满足边值条件, 取

$$
\omega(x)=x(1-x),
$$

将 $u_n(x)$ 表成

$$
u_n(x)=\sum_{i=1}^{n}c_i\varphi_i(x)=x(1-x)\ (c_1+c_2x+\cdots+c_nx^{n-1}).
$$

先令 $n=1$, 则 $u_1=c_1x(1-x)$. 由 (5.5.7) $(n=1), c_1$ 满足方程

$$
c_1\int_0^1({\varphi''}_1+\varphi_1)\varphi_1\mathrm{d}x=-\int_0^1 x^2(1-x)\mathrm{d}x.
$$

经计算, 得

$$
-\frac{3}{10}c_1=-\frac{1}{12}, \quad c_1=\frac{5}{18}, \quad u_1=\frac{5}{18}x(1-x).
$$

再令 $n=2$. 以 $u_2=c_1\varphi_1+c_2\varphi_2, f=-x$ 代到 (5.5.7), 经简单计算, 得 Ritz–Galerkin 方程:

$$
\begin{cases} -\dfrac{3}{10}c_1-\dfrac{3}{20}c_2=-\dfrac{1}{12}, \\ -\dfrac{3}{20}c_1-\dfrac{13}{105}c_2=-\dfrac{1}{20}. \end{cases}
$$

解之, 得 $c_1=\dfrac{71}{369}, c_2=\dfrac{7}{41}$, 于是

$$
u_2=x(1-x)\left(\frac{71}{369}+\frac{7}{41}x\right).
$$

边值问题 (5.5.16) 的精确解为

$$u_* = \frac{\sin x}{\sin 1} - x.$$

表 5.1 列出 $u_1(x), u_2(x), u_*(x)$ 于 $x = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ 的函数值.

表 5.1

x	$u_*(x)$	$u_1(x)$	$u_2(x)$
$\frac{1}{4}$	0.044	0.052	0.044
$\frac{1}{2}$	0.070	0.069	0.069
$\frac{3}{4}$	0.060	0.052	0.060

上述例子是简单的. 实际应用中的问题要复杂得多. 例如基函数的选取, 它必须满足本质边值条件. 在有限元方法出现以前, 通常选代数或三角多项式为基函数, 除特别规则的区域外, 要它们满足边值条件是困难的.

下一节我们将对规则区域和周期边值条件介绍一类有效的谱方法和拟谱法. 第六章介绍的有限元法, 提供了系统构造基函数或子空间的方法, 可用于求解复杂的边值问题.

习 题

用 Ritz–Galerkin 方法求边值问题

$$\begin{cases} u'' + u = x^2, \quad 0 < x < 1, \\ u(0) = 0, \quad u(1) = 1 \end{cases}$$

的第 n 次近似 $u_n(x)$, 基函数为 $\varphi_i(x) = \sin(i\pi x), i = 1, 2, \cdots, n$.

§6 谱 方 法

本节针对规则区域, 例如一维区间, 二维矩形以及三维长方体等乘积型区域和周期边值条件, 介绍 Fourier 谱方法, 这是经典 Ritz–Galerkin 法常用的一种方法. 由于该方法的计算量大, 且要求基函数满足边值条件, 所以在应用中受到很大限制. 1965 年, 出现了计算离散 Fourier 变换的快速算法 – FFT 算法 (参看 [5]), 这不仅给 Fourier 谱方法提供了快速发展的机遇, 而且还将它推广到关于一般正交多项式展开的谱方法 (参看 [13]). 作为模型, 我们考虑两点边值问题, 推广到高维乘积型区域边值问题并不困难.

6.1　三角函数逼近

现在假设 $H^m(0,2\pi)$ 是定义在 $(0,2\pi)$ 取复值的 Sobolev 空间, $H_p^m(0,2\pi)$ 是 $H^m(0,2\pi)$ 中以 2π 为周期的函数组成的 Sobolev 子空间:

$$H_p^m(0,2\pi) = \{f \mid f \in H^m(0,2\pi),\ f(x+2\pi) = f(x)\},$$

其内积和范数分别为

$$(f,g)_m = \int_0^{2\pi} \sum_{s=0}^{m} f^{(s)} \overline{g}^{(s)} \mathrm{d}x$$

和

$$\|f\|_m = \sqrt{(f,f)_m}.$$

往后以 (f,g) 和 $\|f\|$ 分别表示 $(f,g)_0$, 和 $\|f\|_0$.

设 $f \in H^m(0,2\pi)$, 将 $f(x)$ 展成 Fourier 级数:

$$f(x) = \sum_{k=-\infty}^{\infty} \hat{f}(k) \mathrm{e}^{\mathrm{i}kx}, \tag{5.6.1}$$

其中

$$\hat{f}(k) = \frac{1}{2\pi} \int_0^{2\pi} f(x) \mathrm{e}^{-\mathrm{i}kx} \mathrm{d}x = \frac{1}{2\pi}(f, \mathrm{e}^{\mathrm{i}kx})$$

为 f 的 Fourier 系数. 对 (5.6.1) 逐项微商 (求广义导数), 得

$$\frac{\mathrm{d}^s f(x)}{\mathrm{d}x^s} = \sum_{k=-\infty}^{\infty} \hat{f}(k)(\mathrm{i}k)^s \mathrm{e}^{\mathrm{i}kx}, \quad 0 \leqslant s \leqslant m. \tag{5.6.2}$$

由三角函数系 $\{\mathrm{e}^{\mathrm{i}kx}\}_{k=-\infty}^{\infty}$ 的正交性

$$(\mathrm{e}^{\mathrm{i}jx}, \mathrm{e}^{\mathrm{i}kx}) = \int_0^{2\pi} \mathrm{e}^{\mathrm{i}(j-k)x} \mathrm{d}x = \begin{cases} 0, & k \neq j, \\ 2\pi, & k = j, \end{cases} \tag{5.6.3}$$

可得

$$\begin{aligned} \left\| \frac{\mathrm{d}^s f}{\mathrm{d}x^s} \right\|^2 &= \int_0^{2\pi} \left(\sum_{j=-\infty}^{\infty} \hat{f}(j)(\mathrm{i}j)^s \mathrm{e}^{\mathrm{i}jx} \right) \left(\sum_{k=-\infty}^{\infty} \overline{\hat{f}}(k)(-\mathrm{i}k)^s \mathrm{e}^{-\mathrm{i}kx} \right) \mathrm{d}x \\ &= 2\pi \sum_{k=-\infty}^{\infty} k^{2s} |\hat{f}(k)|^2. \end{aligned} \tag{5.6.4}$$

取无穷级数 (5.6.1) 的 $2N+1$ 项和:

$$f_N(x) = \sum_{k=-N}^{N} \hat{f}(x) \mathrm{e}^{\mathrm{i}kx}, \tag{5.6.5}$$

我们自然关心 $f_N(x)$ 对 $f(x)$ 的逼近性.

引理 6.1 设 $f(x) \in H^m(0,2\pi)$, 则对 $s: 0 \leqslant s \leqslant m$, 有估计

$$\|f(x) - f_N(x)\|_s \leqslant CN^{s-m}\|f\|_m,$$

其中 C 是与 f, N 无关的常数.

证明 由 (5.6.1) 和 (5.6.5),

$$\frac{\mathrm{d}^j(f(x)-f_N(x))}{\mathrm{d}x^j} = \frac{\mathrm{d}^j}{\mathrm{d}x^j}\sum_{|n|>N}\hat{f}(n)\mathrm{e}^{\mathrm{i}nx} = \sum_{|n|>N}\hat{f}(n)(\mathrm{i}n)^j\mathrm{e}^{\mathrm{i}nx}.$$

对 $0 \leqslant j \leqslant s \leqslant m$, 由 (5.6.4) 有

$$\begin{aligned}\left\|\frac{\mathrm{d}^j(f(x)-f_N(x))}{\mathrm{d}x^j}\right\|^2 &= 2\pi\sum_{n>N}|\hat{f}(n)|^2 n^{2j} = 2\pi\sum_{n>N}\left|\hat{f}(n)\right|^2 n^{2(j-m)}n^{2m}\\ &\leqslant 2\pi N^{2(j-m)}\sum_{|n|>N}|\hat{f}(n)|^2 n^{2m} \leqslant N^{2(j-m)}\left\|\frac{\mathrm{d}^m f}{\mathrm{d}x^m}\right\|^2.\end{aligned}$$

关于 $j: 0 \leqslant j \leqslant s$ 取和, 两边开方, 即知结论成立.

记 $U_N = \mathrm{span}\{\mathrm{e}^{\mathrm{i}kx}\}_{k=-N}^{N}, P_N$ 是由 $H = H^0$ 到 U_N 的 L^2 投影算子, 即对 $\forall f \in H$, 有唯一 $P_N f \in U_N$, 使

$$(P_N f, v) = (f, v), \quad \forall v \in U_N.$$

则由正交性 (5.6.3) 知, $f_N(x)$ 就是 $f(x)$ 从 $H(0,2\pi)$ 到 U_N 的 L^2 投影, 记为 $f_N = P_N f$.

为应用快速 Fourier 变换 (FFT) 算法, 最好取 $N2^m : U_N = \mathrm{span}\{\mathrm{e}^{\mathrm{i}kx}\}_{k=-N}^{N-1}$. 下面考虑 $f(x)$ 在 U_N 的插值逼近. 在 $[0,2\pi]$ 中引入 $2N$ 个等距节点 $x_m = \dfrac{m\pi}{N}, m = 0, 1, \cdots, 2N$. 利用三角函数的性质可以证明

$$\frac{1}{2N}\sum_{k=-N}^{N-1}\mathrm{e}^{\mathrm{i}kx_m} = \begin{cases}1, m = 0,\\ 0, m \neq 0.\end{cases} \tag{5.6.6}$$

令

$$l_m(x) = \frac{1}{2N}\sum_{k=-N}^{N-1}\mathrm{e}^{\mathrm{i}k(x-x_m)}. \tag{5.6.7}$$

由 (5.6.6),

$$l_m(x_n) = \delta_{mn} = \begin{cases}1, & m = n,\\ 0, & m \neq n.\end{cases}$$

可见 $l_m(x)$ 可作为 Lagrange 插值基函数. 这样, $f(x)$ 在 U_N 中以 $\{x_j\}$ 为插值节点的插值多项式为

$$(5.6.8)\qquad \begin{aligned} I_N f(x) &= \sum_{m=1}^{2N} f(x_m) l_m(x) = \sum_{m=1}^{2N} f(x_m) \frac{1}{2N} \sum_{k=-N}^{N-1} \mathrm{e}^{\mathrm{i}k(x-x_m)} \\ &= \sum_{k=-N}^{N-1} \mathrm{e}^{\mathrm{i}kx} \frac{1}{2N} \sum_{m=1}^{2N} f(x_m) \mathrm{e}^{-\mathrm{i}kx_m}. \end{aligned}$$

定义离散内积和范数:

$$(5.6.9)\qquad (u,v)_N = \frac{\pi}{N} \sum_{m=1}^{2N} u(x_m)\overline{v}(x_m), ||u||_N^2 = (u,u)_N.$$

则

$$(5.6.10)\qquad I_N f(x) = \sum_{k=-N}^{N-1} \tilde{f}(k) \mathrm{e}^{\mathrm{i}kx},$$

其中

$$(5.6.11)\qquad \tilde{f}(k) = \frac{1}{2\pi}(f, \mathrm{e}^{\mathrm{i}kx})_N$$

是离散 Fourier 系数.

6.2 Fourier 谱方法

现在举例说明谱方法的应用.

考虑求解周期边值问题:

$$(5.6.12)_1\qquad Lu \equiv -u'' + \lambda u = f(x), \quad x \in (0, 2\pi),$$

$$(5.6.12)_2\qquad u(0) = u(2\pi) = 0,$$

其中 $\lambda > 0$ 是常数, $f(x)$ 为 2π- 周期函数. 引进空间

$$(5.6.13)\qquad U = \left\{ u \in H^1(0, 2\pi), u(0) = u(2\pi) = 0 \right\},$$

用 v 的复共轭函数 $\overline{v}(x) \in U$ 乘方程 $(5.6.12)_1$ 两端, 在 $[0, 2\pi]$ 上积分, 并施行分部积分. 利用周期性边值条件 $(5.6.12)_2$, 可得

$$\begin{aligned} \int_0^{2\pi} Lu \cdot \overline{v} \mathrm{d}x &= \int_0^{2\pi} u'\overline{v}' \mathrm{d}x - u'\overline{v}\big|_0^{2\pi} + \int_0^{2\pi} \lambda u \overline{v} \mathrm{d}x \\ &= \int_0^{2\pi} u'\overline{v'} \mathrm{d}x + \int_0^{2\pi} \lambda u \overline{v} \mathrm{d}x = \int_0^{2\pi} f \overline{v} \mathrm{d}x. \end{aligned}$$

令

$$a(u,v)=\int_0^{2\pi}(u'\overline{v}'+\lambda u\overline{v})\mathrm{d}x,\quad (f,v)=\int_0^{2\pi}f\overline{v}\mathrm{d}x.$$

于是问题 (5.6.12) 的变分形式为: 求 $u\in U$, 使得

(5.6.14) $$a(u,v)=(f,v),\quad \forall v\in U.$$

在 Galerkin 法中, 取子空间 $U_n=\mathrm{span}\{\mathrm{e}^{ikx}\}_{k=-N}^{N}$, 就导出所谓 Fourier 谱方法.

现在介绍解边值问题 (5.6.12) 的谱方法. 取基函数 $\varphi_k=\mathrm{e}^{ikx}-1,k=\pm1,\pm2,\cdots,\pm N,U_N=\mathrm{span}\{\phi_k;k=\pm1,\pm2,\cdots,\pm N\}$, 则谱方法为: 求

(5.6.15)
$$u_N=\sum_{\substack{k=-N\\k\neq0}}^{N}c_k\phi_k=\sum_{\substack{k=-N\\k\neq0}}^{N}c_k\left(\mathrm{e}^{ikx}-1\right)=\sum_{k=-N}^{N}c_k\mathrm{e}^{ikx},$$
$$c_0=-\sum_{k=-N,k\neq0}^{N}c_k,$$

满足

(5.6.16)
$$\begin{aligned}a(u_N,\varphi_j)&=\sum_{k=-N}^{N}c_ka(\mathrm{e}^{ikx},\varphi_j)\\&=\sum_{k=-N}^{N}c_ka(\mathrm{e}^{ikx},\mathrm{e}^{ijx})-\sum_{k=-N}^{N}c_ka(\mathrm{e}^{ikx},1),\quad j=\pm1,\pm2,\cdots,\pm N,\\&=(f,\varphi_j)\end{aligned}$$

和

(5.6.17) $$\sum_{k=-N}^{N}c_k=0.$$

由正交性 (5.6.3) 得

(5.6.18)
$$a(\mathrm{e}^{ikx},\mathrm{e}^{ijx})=\int_0^{2\pi}(-(\mathrm{i}k)(\mathrm{i}j)\mathrm{e}^{ikx}\mathrm{e}^{-\mathrm{i}jx}+\lambda\mathrm{e}^{ikx}\mathrm{e}^{-\mathrm{i}jx})\mathrm{d}x=\begin{cases}0, & k\neq j,\\2\pi(\lambda+j^2), & k=j,\end{cases}$$

特别取 $k=j=0$, 得

$$a(1,1)=2\pi\lambda.$$

又

(5.6.19) $$(f,\varphi_j)=\int_0^{2\pi}f(x)\mathrm{e}^{-\mathrm{i}jx}\mathrm{d}x-\int_0^{2\pi}f(x)\mathrm{d}x.$$

于是方程组 (5.6.16) 化为

$$2\pi(\lambda+j^2)c_j-2\pi\lambda c_0=(f,\varphi_j),\quad j=\pm1,\pm2,\cdots,\pm N.$$

与 (5.6.17) 联立得方程组

$$\begin{bmatrix}1 & & & & & a_{-1}\\ & 1 & & & & a_1\\ & & \ddots & & & \vdots\\ & & & 1 & & a_{-N}\\ & & & & 1 & a_N\\ -1 & -1 & -1 & -1 & -1 & -1\end{bmatrix}\begin{bmatrix}c_{-1}\\ c_1\\ \vdots\\ c_{-N}\\ c_N\\ c_0\end{bmatrix}=\begin{bmatrix}b_{-1}\\ b_1\\ \vdots\\ b_{-N}\\ b_N\\ 0\end{bmatrix},\tag{5.6.20}$$

其中

$$a_j=-\frac{\lambda}{\lambda+j^2},\quad b_j=\frac{(f,\varphi_j)}{2\pi(\lambda+j^2)},\quad j=\pm1,\pm2,\cdots,\pm N.\tag{5.6.21}$$

用消元法解出系数 c_j 即得近似解.

由定理 5.1 和引理 6.1 得

$$\|u-u_N\|_1\leqslant C\inf_{v_n\in U_n}\|u-v_n\|_1\leqslant CN^{-(s-1)}\|u\|_s.$$

由上式可看出, $u(x)$ 的光滑性越好, $u_N(x)$ 收敛得越快. 特别, 若 $u(x)$ 是无穷次可微的周期函数, 则 $u_N(x)$ 收敛于 $u(x)$ 的速度快于 $\dfrac{1}{N}$ 的任何有限次幂. 所以也说谱方法具有 "指数收敛性".

注 Fourier 谱方法要求解具有周期性. 对非周期情形, 可作周期性扩张, 但在边界点会出现间断, 将产生不应有的振荡. 为此, 人们研究用 Chebyshev, Legendre 等多项式作为逼近工具, 并简称相应的 Galerkin 方法为谱方法 (参看 [13]).

数值例子 用谱方法求解

$$\begin{aligned}&-u''+u=2x\sin x-2\cos x,\quad x\in(0,2\pi)\\ &u(0)=u(2\pi)=0\\ &(\text{精确解 } u=x\sin x).\end{aligned}$$

令 $f(x)=2x\sin x-2\cos x$, 谱方法右端

$$\begin{aligned}(f,\phi_j)&=\int_0^{2\pi}f\overline{\phi}_j\mathrm{d}x=\int_0^{2\pi}f(x)\mathrm{e}^{-\mathrm{i}jx}\mathrm{d}x-\int_0^{2\pi}f(x)\mathrm{d}x\\ &=\int_0^{2\pi}f(x)\cos jx\mathrm{d}x-\int_0^{2\pi}f(x)\mathrm{d}x-\mathrm{i}\int_0^{2\pi}f(x)\sin jx\mathrm{d}x.\end{aligned}\tag{5.6.22}$$

经计算得

$$(f,\phi_j) = \begin{cases} 4\pi\dfrac{j^2}{j^2-1}, & j\neq\pm1, \\ \pi-2i\pi^2, & j=1, \\ \pi+2i\pi^2, & j=-1. \end{cases} \tag{5.6.23}$$

另一方面, 由 (5.6.21) 有 (取 $\lambda=1$)

$$b_j=(f,\varphi_j)/2\pi(1+j^2)=\begin{cases} -\dfrac{2j^2}{(1-j^4)}, & j\neq\pm1, \\ \dfrac{1}{4}-i\dfrac{\pi}{2}, & j=1, \\ \dfrac{1}{4}+i\dfrac{\pi}{2}, & j=-1. \end{cases} \tag{5.6.24}$$

$$a_j=\frac{-1}{1+j^2}, \quad j=\pm1,\pm2,\cdots,\pm N. \tag{5.6.25}$$

最后解方程组 (5.6.20). 用消元法得

$$\begin{bmatrix} 1 & & & & & a_{-1} \\ & 1 & & & & a_1 \\ & & \ddots & & & \vdots \\ & & & 1 & & a_{-N} \\ & & & & 1 & a_N \\ & & & & & d_0 \end{bmatrix} \begin{bmatrix} c_{-1} \\ c_1 \\ \vdots \\ c_{-N} \\ c_N \\ c_0 \end{bmatrix} = \begin{bmatrix} b_{-1} \\ b_1 \\ \vdots \\ b_{-N} \\ b_N \\ b_0 \end{bmatrix}, \tag{5.6.26}$$

其中

$$b_0=\sum_{k=-N,k\neq0}^{N} b_k=-4\sum_{j=2}^{N}\frac{j^2}{(1-j^4)}+\frac{1}{2}, \tag{5.6.27}$$

$$d_0=-1+\sum_{k=-N,k\neq0}^{N} a_k=-\left(1+2\sum_{k=1}^{N}\frac{1}{1+k^2}\right). \tag{5.6.28}$$

由 (5.6.26) 回代得

$$c_0=\frac{b_0}{d_0}, \quad c_j=b_j-a_jc_0. \tag{5.6.29}$$

从而

$$\begin{aligned} u_N &= c_0+\sum_{j=-N,j\neq0}^{N} c_j e^{ijx} = c_0+\sum_{j=-N,j\neq0}^{N} b_j e^{ijx} - c_0\sum_{j=-N,j\neq0}^{N} a_j e^{ijx} \\ &= c_0\left(1-\sum_{j=-N,j\neq0}^{N} a_j e^{ijx}\right)+\sum_{j=-N,j\neq0}^{N} b_j e^{ijx}, \end{aligned} \tag{5.6.30}$$

其中 (利用 (5.6.24) 和 (5.6.25))

$$(5.6.31)\qquad \sum_{j=-N,j\neq 0}^{N} a_j e^{ijx} = \sum_{j=1}^{N} \frac{-2}{1+j^2}\cos jx,$$

$$(5.6.32)\qquad \sum_{j=-N,j\neq 0}^{N} b_j e^{ijx} = -4\sum_{j=2}^{N} \frac{j^2}{(1-j^4)}\cos jx + \frac{1}{2}\cos x + \pi \sin x.$$

将 (5.6.27)—(5.6.29) 和 (5.6.31)—(5.6.32) 代入 (5.6.30), 则得近似解:

$$(5.6.33)\qquad \begin{aligned} u_N &= c_0\left(1 - \sum_{j=-N,j\neq 0}^{N} a_j e^{ijx}\right) + \sum_{j=-N,j\neq 0}^{N} b_j e^{ijx} \\ &= \left\{-\left(1 + 2\sum_{k=1}^{N}\frac{1}{1+k^2}\right)^{-1}\left[-4\sum_{j=2}^{N}\frac{j^2}{(1-j^4)} + \frac{1}{2}\right]\right\}\cdot \\ &\quad \left\{1 + 2\sum_{j=1}^{N}\frac{1}{1+j^2}\cos jx\right\} - \\ &\quad 4\sum_{j=2}^{N}\frac{j^2}{(1-j^4)}\cos jx + \frac{1}{2}\cos x + \pi\sin x\ . \end{aligned}$$

计算结果如图 6.1, 图 6.2, 可见谱方法的精度是很高的.

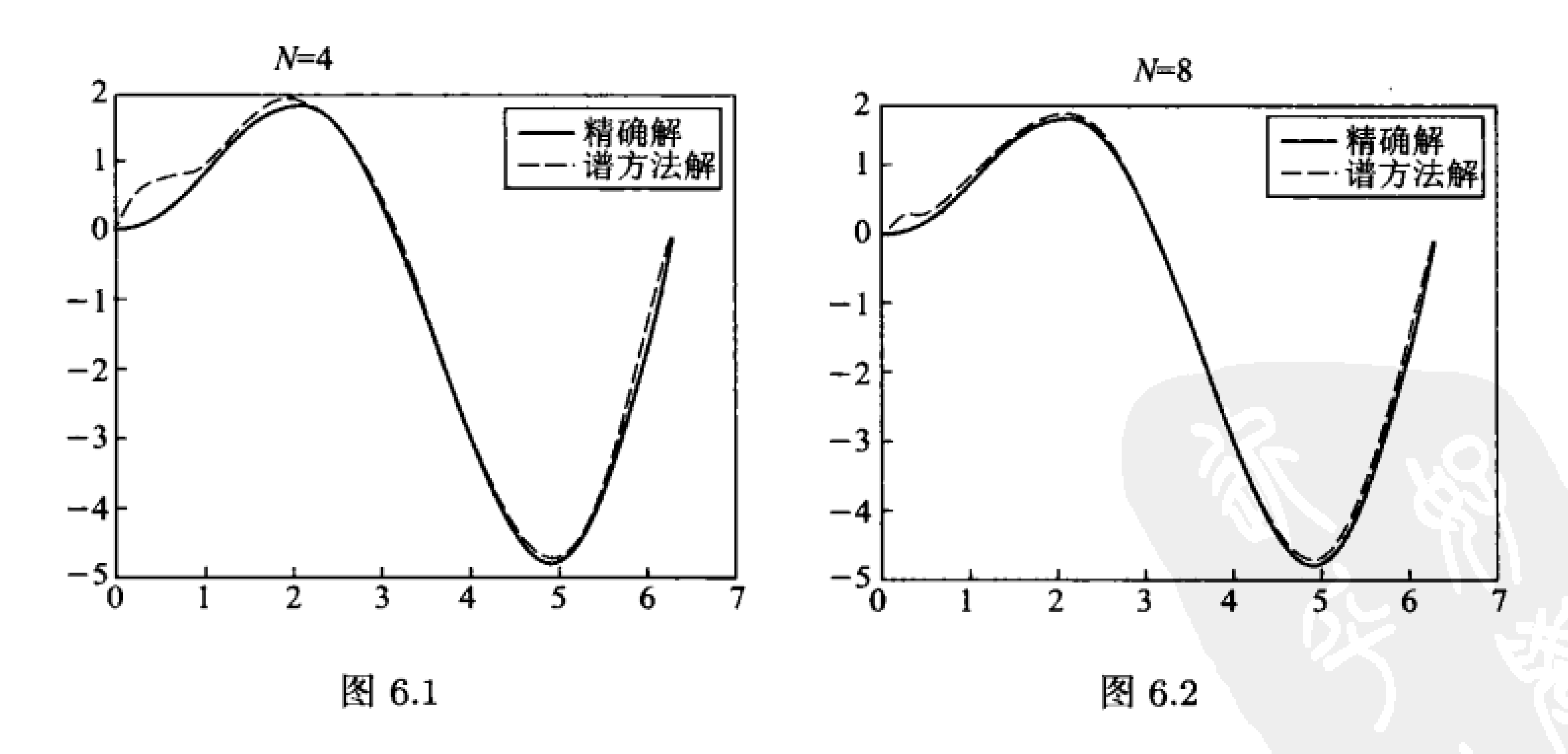


图 6.1　　图 6.2

6.3　拟谱方法 (配置法)

谱方法要计算许多诸如 $a(e^{ikx}, e^{ijx})$ 的内积, 对变系数方程, 计算量较大, 有时要用数值积分公式. 现在采用配置法, 称为拟谱方法, 可明显减少计算量.

设边值问题为求 $2\pi-$ 周期函数 u, 满足

$$Lu = f. \tag{5.6.34}$$

拟谱方法是选定节点组 $\{x_j\}$, 求 $u_N \in U_N$, 使得 u_N 在 $\{x_j\}$ 上满足方程 (5.6.34).

仍以边值问题 (5.6.12) 为例介绍拟谱方法的应用.

取基函数 $\varphi_k = \mathrm{e}^{\mathrm{i}kx} - 1, k = \pm 1, \pm 2, \cdots, \pm N, U_N = \mathrm{span}\{\phi_k; k = \pm 1, \pm 2, \cdots, \pm N\}$, 求

$$u_N = \sum_{\substack{k=-N \\ k\neq 0}}^{N} c_k \phi_k = \sum_{\substack{k=-N \\ k\neq 0}}^{N} c_k \left(\mathrm{e}^{\mathrm{i}kx} - 1\right) = \sum_{k=-N}^{N} c_k \mathrm{e}^{\mathrm{i}kx},$$

$$c_0 = -\sum_{k=-N, k\neq 0}^{N} c_k. \tag{5.6.35}$$

则

$$\frac{\mathrm{d}u_N}{\mathrm{d}x} = \sum_{k=-N}^{N} c_k (\mathrm{i}k) \mathrm{e}^{\mathrm{i}kx}, \quad \frac{\mathrm{d}^2 u_N}{\mathrm{d}x^2} = \sum_{k=-N}^{N} c_k (-k^2) \mathrm{e}^{\mathrm{i}kx}.$$

由于 $\mathrm{e}^{\mathrm{i}kx_0} = \mathrm{e}^{\mathrm{i}kx_{2N}}$, 故可设配置点为 $x_m = \dfrac{m\pi}{N}, m = 1, \cdots, 2N$, 令 u_N 在 $2N$ 个配置点 x_m 上满足方程 (5.6.12), 得到

$$\sum_{k=-N}^{N} c_k k^2 \mathrm{e}^{\mathrm{i}kx_m} + \lambda \sum_{k=-N}^{N} c_k \mathrm{e}^{\mathrm{i}kx_m} = f(x_m), \quad m = 1, \cdots, 2N. \tag{5.6.36}$$

两端乘以 $\mathrm{e}^{-\mathrm{i}jx_m}$, 并关于 $m = 1, \cdots, 2N$ 求和, 则左端为

$$\begin{aligned}
&\sum_{m=1}^{2N} \sum_{k=-N}^{N} c_k k^2 \mathrm{e}^{\mathrm{i}kx_m} \mathrm{e}^{-\mathrm{i}jx_m} + \lambda \sum_{m=1}^{2N} \sum_{k=-N}^{N} c_k \mathrm{e}^{\mathrm{i}kx_m} \mathrm{e}^{-\mathrm{i}jx_m} \\
= &\sum_{k=-N}^{N} c_k k^2 \sum_{m=1}^{2N} \mathrm{e}^{\mathrm{i}kx_m} \mathrm{e}^{-\mathrm{i}jx_m} + \lambda \sum_{k=-N}^{N} c_k \sum_{m=1}^{2N} \mathrm{e}^{\mathrm{i}kx_m} \mathrm{e}^{-\mathrm{i}jx_m} \\
= &\sum_{k=-N}^{N} c_k k^2 \frac{N}{\pi} (\mathrm{e}^{\mathrm{i}kx}, \mathrm{e}^{\mathrm{i}jx})_N + \lambda \sum_{k=-N}^{N} c_k \frac{N}{\pi} (\mathrm{e}^{\mathrm{i}kx}, \mathrm{e}^{\mathrm{i}jx})_N,
\end{aligned}$$

右端为

$$\sum_{m=1}^{2N} f(x_m)\, \mathrm{e}^{-\mathrm{i}jx_m} = \frac{N}{\pi} (f, \mathrm{e}^{\mathrm{i}jx})_N,$$

其中

$$(u, v)_N = \frac{\pi}{N} \sum_{m=1}^{2N} u(x_m) \overline{v}(x_m).$$

于是方程组 (5.6.36) 化为

$$(5.6.37)\qquad \sum_{k=-N}^{N} c_k(k^2+\lambda)(\mathrm{e}^{\mathrm{i}kx},\mathrm{e}^{\mathrm{i}jx})_N=(f,\mathrm{e}^{\mathrm{i}jx})_N,\quad j=-N,-N+1,\cdots,N$$

与

$$c_0=-\sum_{k=-N,k\neq 0}^{N} c_k.$$

联立求出 $c_k, k=0,\pm1,\cdots,\pm N$, 代到 (5.6.35) 即得 u_N.

(5.6.37) 是 Fourier 谱方法中方程 (5.6.16) 的离散形式. 计算形如 $(\mathrm{e}^{\mathrm{i}kx},\mathrm{e}^{\mathrm{i}jx})_N,(f,\mathrm{e}^{\mathrm{i}jx})_N$ 的离散内积时, 可采用快速 Fourier 变换算法 (FFT) (参考 [5]). FFT 可将要计算的复数运算由 $O(N^2)$ 个, 减少到 $O(N\cdot\log_2 N)$ 个. 可以证明, 拟谱方法和谱方法有同样的收敛阶. 对于具复杂系数的方程, 特别是非线性问题, 拟谱方法更为实用 (参看 [13]).

第六章 Galerkin 有限元法

有限元法, 实质上就是 Ritz–Galerkin 法, 它和传统的 Ritz–Galerkin 法的主要区别在于, 它用样条函数方法提供了一种选取 "局部基函数" 或 "分片多项式空间" 的技术, 从而在很大程度上克服了 Ritz–Galerkin 法选取基函数的困难. 有限元法首先成功地用于结构力学和固体力学, 以后又用于流体力学、物理学和其他工程科学. 现在, 有限元法和差分法一样, 已成为求解偏微分方程, 特别是椭圆型偏微分方程的一种有效数值方法.

有限元法的基本问题可归纳为:

(1) 把问题转化成变分形式.

(2) 选定单元的形状, 对求解域作剖分.

一维情形的单元是小区间. 二维情形的重要单元有两种: 四边形 (矩形、任意凸四边形) 和三角形. 三维单元就更复杂多样了. 本书只讨论一维、二维单元.

(3) 构造基函数或单元形状函数.

(4) 形成有限元方程 (Ritz–Galerkin 方程).

(5) 提供有限元方程的有效解法.

(6) 收敛性及误差估计.

第一个问题已在第五章讲过了, 第五个问题已在第二章最后三节讨论. 本章 §7 还将介绍多重网格计算技术. 为了便于读者理解, 我们先讲一维域上两点边值问题的有限元法 (§1,§2), 然后推广到平面域上二阶椭圆边值问题 (§3—§6). 最后在 §8, 介绍如何将有限元法推广到初边值问题, 包括抛物方程和二阶双曲方程. 我们仅就两点边值问题讨论线性有限元法的收敛性和误差估计. 有限元法的一般理论已很完善, 有兴趣的读者可参看 [10]、[21].

§1 两点边值问题的有限元法

考虑两点边值问题

$$\begin{cases} Lu = -\dfrac{\mathrm{d}}{\mathrm{d}x}\left(p\dfrac{\mathrm{d}u}{\mathrm{d}x}\right) + qu = f, \quad a < x < b, \\ u(a) = 0, \quad u'(b) = 0, \end{cases} \tag{6.1.1}$$

其中 $p \in C^1(I)$, $p \geqslant p_0 > 0$, $q \in C(I)$, $q \geqslant 0$, $f \in C(I)$, $I = [a, b]$. 我们将从 Ritz 法和 Galerkin 法两种观点出发, 导出解 (6.1.1) 的有限元法.

1.1 从 Ritz 法出发

首先对求解区间作网格剖分, 节点为

$$a = x_0 < x_1 < \cdots < x_n = b,$$

相邻节点 x_{i-1}, x_i 之间的小区间 $I_i = [x_{i-1}, x_i]$ 称为单元 (第 i 号单元), 其长度为 $h_i = x_i - x_{i-1}$.

其次, 在 Sobolev 空间 $H_E^1(I)$ (见第五章 §2) 内取子空间 U_h (下标 $h = \max h_i$), 它的元素 $u_h(x)$ 在每一单元上是次数不超过某一正整数 m 的多项式, 在全区间 $[a, b]$ 上属于函数空间 $H_E^1(I)$, 就是说, $u_h(x) \in H^1(I)$ 且 $u_h(a) = 0$. 显然这是有限维空间, 称 U_h 为**试探函数空间**, $u_h \in U_h$ 为**试探函数** (trial function).

最简单的试探函数空间 U_h 是由分段线性函数组成的, 它由节点上的一组值

$$u_0 = 0, u_1, u_2, \cdots, u_n$$

按线性插值公式

$$u_h(x) = \frac{x_i - x}{h_i} u_{i-1} + \frac{x - x_{i-1}}{h_i} u_i, \quad x \in I_i, \quad i = 1, 2, \cdots, n \tag{6.1.2}$$

确定, 称为线性有限元 (一次有限元) 空间. 其几何形状如图 1.1. 称 (6.1.2) 为单元形状函数, 它是试探函数 在单元 I_i 上的限制 (restriction).

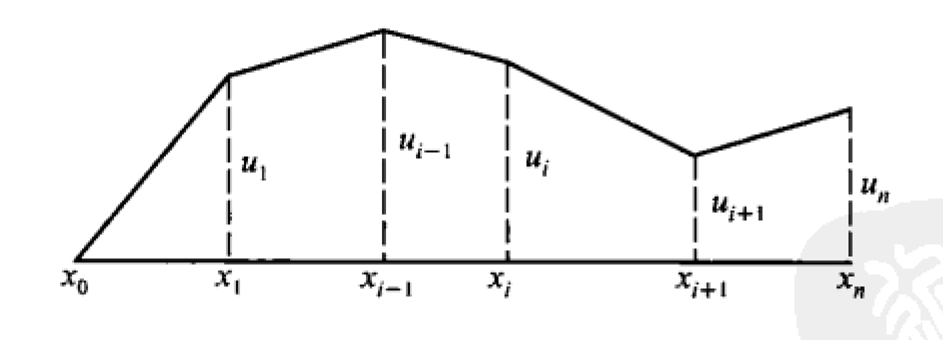


图 1.1

为使按段插值标准化, 通常用仿射变换

$$\xi = F_i(x) = \frac{x - x_{i-1}}{h_i} \tag{6.1.3}$$

把 I_i 轴变到 ξ 轴上的参考单元 $[0, 1]$. 令

$$N_0(\xi) = 1 - \xi, \quad N_1(\xi) = \xi,$$

则

$$u_h(x) = N_0(\xi)u_{i-1} + N_1(\xi)u_i, \quad x \in I_i. \tag{6.1.4}$$

因为 u_h 的自由度是 n, 故 U_h 是 n 维线性空间.

将 (6.1.2) 代入泛函数

$$J(u) = \frac{1}{2}\int_a^b (pu'^2 + qu^2 - 2fu)\mathrm{d}x,$$

由第五章 (5.3.5) 式, 得

$$\begin{aligned} J(u_h) &= \frac{1}{2}\int_a^b [pu'^2_h + qu_h^2 - 2fu_h]\mathrm{d}x \\ &= \frac{1}{2}\sum_{i=1}^n \int_{I_i} [pu'^2_h + qu_h^2]\mathrm{d}x - \sum_{i=1}^n \int_{I_i} fu_h \mathrm{d}x. \end{aligned}$$

利用变换 (6.1.3), 则

$$\begin{aligned} J(u_h) = \frac{1}{2}\sum_{i=1}^n \int_0^1 \bigg[& p(x_{i-1} + h_i\xi)\frac{(u_i - u_{i-1})^2}{h_i} + \\ & h_i q(x_{i-1} + h_i\xi)(N_0(\xi)u_{i-1} + N_1(\xi)u_i)^2\bigg]\mathrm{d}\xi - \\ & \sum_{i=1}^n h_i \int_0^1 f(x_{i-1} + h_i\xi)(N_0(\xi)u_{i-1} + N_1(\xi)u_i)\mathrm{d}\xi. \end{aligned} \tag{6.1.5}$$

令

$$\frac{\partial J(u_h)}{\partial u_j} = a_{j-1,j}u_{j-1} + a_{jj}u_j + a_{j+1,j}u_{j+1} - b_j = 0, \tag{6.1.6}$$

其中

$$\begin{cases} a_{j-1,j} = \displaystyle\int_0^1 [-h_j^{-1}p(x_{j-1} + h_j\xi) + h_j q(x_{j-1} + h_j\xi)\xi(1-\xi)]\mathrm{d}\xi, \\ a_{j+1,j} = \displaystyle\int_0^1 [-h_{j+1}^{-1}p(x_j + h_{j+1}\xi) + h_{j+1} q(x_j + h_{j+1}\xi)\xi(1-\xi)]\mathrm{d}\xi, \\ a_{jj} = \displaystyle\int_0^1 [h_j^{-1}p(x_{j-1} + h_j\xi) + h_j q(x_{j-1} + h_j\xi)\xi^2]\mathrm{d}\xi + \\ \displaystyle\int_0^1 [h_{j+1}^{-1}p(x_j + h_{j+1}\xi) + h_{j+1} q(x_j + h_{j+1}\xi)(1-\xi)^2]\mathrm{d}\xi, \\ b_j = h_j \displaystyle\int_0^1 f(x_{j-1} + h_j\xi)\xi\mathrm{d}\xi + h_{j+1}\int_0^1 f(x_j + h_{j+1}\xi)(1-\xi)\mathrm{d}\xi, \end{cases} \tag{6.1.6$'$}$$

就得到 $u_1, u_2, \cdots, u_n$ 的线性代数方程组, 称之为**有限元方程**.

在工程计算中, 并不是直接由 (6.1.6), (6.1.6)′ 形成有限元方程, 而是分析每一单元的局部二次形及单元矩阵, 力学上称为**单元刚度矩阵**; 再由单元刚度矩阵形成总矩阵, 称为**总刚度矩阵**. 刚度矩阵分析法比较灵活, 程序也易实现. 考察 (6.1.5) 右端第一个求和号内第 i 项 (对应第 i 个单元), 它是 u_{i-1}, u_i 的二次形, 可写成形式

$$(\boldsymbol{u}^{(i)})^{\mathrm{T}}\boldsymbol{K}^{(i)}\boldsymbol{u}^{(i)}, \tag{6.1.7}$$

这里 $\boldsymbol{u}^{(i)}=(u_{i-1},u_i)^{\mathrm{T}}$,

$$\boldsymbol{K}^{(i)}=\begin{pmatrix} a_{i-1,i-1}^{(i)} & a_{i-1,i}^{(i)} \\ a_{i,i-1}^{(i)} & a_{ii}^{(i)} \end{pmatrix} \tag{6.1.8}$$

为单元刚度矩阵,

$$\begin{cases} a_{i-1,i-1}^{(i)}=\displaystyle\int_0^1[h_i^{-1}p(x_{i-1}+h_i\xi)+h_iq(x_{i-1}+h_i\xi)(1-\xi)^2]\mathrm{d}\xi, \\ a_{ii}^{(i)}=\displaystyle\int_0^1[h_i^{-1}p(x_{i-1}+h_i\xi)+h_iq(x_{i-1}+h_i\xi)\xi^2]\mathrm{d}\xi, \\ a_{i-1,i}^{(i)}=a_{i,i-1}^{(i)}=\displaystyle\int_0^1[-h_i^{-1}p(x_{i-1}+h_i\xi)+h_iq(x_{i-1}+h_i\xi)\xi(1-\xi)]\mathrm{d}\xi. \end{cases} \tag{6.1.9}$$

现把 $\boldsymbol{K}^{(i)}$ 扩展成 $n\times n$ 矩阵, 第 $i-1$ 行第 i 行和第 $i-1$ 列第 i 列交叉位置的元素就是 (6.1.8) 的四个元素, 其余元素全是零. 记向量 $\boldsymbol{u}=(u_1,u_2,\cdots,u_n)^{\mathrm{T}}$, 则 (6.1.7) 可写成

$$\boldsymbol{u}^{\mathrm{T}}\boldsymbol{K}^{(i)}\boldsymbol{u}.$$

于是 (6.1.5) 右端第一个和式等于

$$\frac{1}{2}\sum_{i=1}^{n}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{K}^{(i)}\boldsymbol{u}=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\left(\sum_{i=1}^{n}\boldsymbol{K}^{(i)}\right)\boldsymbol{u}=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{u},$$

其中

$$\boldsymbol{K}=\sum_{i=1}^{n}\boldsymbol{K}^{(i)} \tag{6.1.10}$$

就是总刚度矩阵.

从 $\boldsymbol{K}^{(i)}$ 形成 $\boldsymbol{K}$ 的步骤是: 第一, 按 (6.1.9) 计算第 i 个单元刚度矩阵的元素 $a_{i-1,i-1}^{(i)},a_{ii}^{(i)},a_{i-1,i}^{(i)}=a_{i,i-1}^{(i)}$, 并标出在总刚度矩阵中的行列号, 如现在的两个下标那样. 第二, 总刚度矩阵 $\boldsymbol{K}$ 的第 i 行第 j 列元素 a_{ij}, 恰好等于所有单元刚度矩阵中具同样行列标号的非零元素的累加和. 实际上, 当 j 不与 i 相邻, 即

$|i-j|>1$ 时 $a_{ij}=0$; 当 j 与 i 相邻, 即 $j=i-1,i,i+1$ 时非零元素只出现在第 i 个节点相邻两单元的刚度矩阵 $\boldsymbol{K}^{(i)}$, $\boldsymbol{K}^{(i+1)}$ 中, 其余单元刚度矩阵不对 a_{ij} 作贡献.

按照上述法则, 不难看出

$$a_{ij}=\begin{cases} a_{i,j-1}^{(i)}, & j=i-1, \\ a_{ii}^{(i)}+a_{ii}^{(i+1)}, & j=i, \\ a_{i,i+1}^{(i+1)}, & j=i+1, \\ 0, & |j-i|\geqslant 2, \end{cases} \tag{6.1.11}$$

因此 $\boldsymbol{K}$ 是三对角矩阵.

用同样方法可以形成 (6.1.5) 右端第二个和式. 令

$$f^{(i)}=(f_{i-1}^{(i)},f_i^{(i)})^{\mathrm{T}},$$

$$\begin{cases} f_{i-1}^{(i)}=h_i\displaystyle\int_0^1 f(x_{i-1}+h_i\xi)(1-\xi)\mathrm{d}\xi, \\ f_i^{(i)}=h_i\displaystyle\int_0^1 f(x_{i-1}+h_i\xi)\xi\mathrm{d}\xi, \end{cases} \tag{6.1.12}$$

以及 $\boldsymbol{b}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}$,

$$b_1=f_1^{(1)}+f_1^{(2)},\quad b_2=f_2^{(2)}+f_2^{(3)},\cdots,\quad b_n=f_n^{(n)}, \tag{6.1.13}$$

则 (6.1.5) 右端第二个和式为

$$\boldsymbol{u}^{\mathrm{T}}\boldsymbol{b}=\sum_{i=1}^{n}(\boldsymbol{u}^{(i)})^{\mathrm{T}}\boldsymbol{f}^{(i)}. \tag{6.1.14}$$

综上所述, (6.1.5) 可写成

$$J(u_h)=\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{u}-\boldsymbol{u}^{\mathrm{T}}\boldsymbol{b}. \tag{6.1.5$'$}$$

因此有限元方程为

$$\boldsymbol{K}\boldsymbol{u}=\boldsymbol{b}. \tag{6.1.15}$$

注 1.1 当第一边值条件 (左边值条件) 非齐次时, 例如 $u(a)=\alpha$, 则需像其他单元一样形成 $I_1=[x_0,x_1]$ 上的单元刚度矩阵. 但形成总刚度矩阵 $\boldsymbol{K}$ 时, 先把 $u_0=\alpha$ 当作未知量, $\boldsymbol{K}$ 扩大成 $(n+1)\times(n+1)$ 矩阵. 然后去掉第一行 (或者

一开始就不计算第一行), 再把第一列的第 j 行元素 a_{j0} 乘以 $(-u_0) = (-\alpha)$, 累加到第 j 个方程的右端后, 再去掉第一列. 最后仍然归结到方程 (6.1.5), 只不过右端向量因第一边值作了修改.

注 1.2 若第二边值条件 (右边值条件) 非齐次, 例如 $u'(b) = \beta$, 则需从下列泛函出发:

$$J(u) = \frac{1}{2}\int_a^b [pu'^2 + qu^2 - 2fu]\mathrm{d}x - \beta p(b)u_n.$$

(参看第五章 §3 习题 1) 它比齐边值多了第二项, 而且是一次项, 不影响总刚度矩阵, 故唯一的改变是第 n 个方程的右端要累加 $\beta p(b)$.

对于第三边值条件

$$u'(b) + \alpha u(b) = \beta,$$

则不但要修改第 n 个方程的右端, 而且总刚度矩阵的第 n 行 n 列元素也要作适当修改, 请读者自己推导.

1.2 从 Galerkin 法出发

现在从 Galerkin 法 (基于虚功原理) 出发推导有限元方程. 为此需要构造试探函数空间 U_h 的一组基底. 我们指出, 同一空间 U_h, 可取各种不同基底, 但并非任一组基底对计算都是可取的. 在单元 I_i, I_{i+1} 考察线性插值公式 (6.1.2) 及 u_i 的系数, 我们自然对每一节点 x_i 构造山形函数

$$\left\{\begin{array}{l} \varphi_i(x) = \begin{cases} 1 + \dfrac{x - x_i}{h_i}, & x_{i-1} \leqslant x \leqslant x_i, \\ 1 - \dfrac{x - x_i}{h_{i+1}}, & x_i \leqslant x \leqslant x_{i+1}, \\ \quad 0, & \text{其他}, \end{cases} \\ \qquad i = 1, 2, \cdots, n-1, \\ \varphi_n(x) = \begin{cases} 1 + \dfrac{x - x_n}{h_n}, & x_{n-1} \leqslant x \leqslant x_n, \\ \quad 0, & \text{其他}, \end{cases} \end{array}\right. \tag{6.1.16}$$

其几何形状如图 1.2.

显然 $\varphi_1(x), \cdots, \varphi_n(x)$ 线性无关, 它们组成 U_h 的基底. 任一 $u_h \in U_h$ 可表成

$$u_h(x) = \sum_{i=1}^{n} u_i\varphi_i(x), \quad u_i = u_h(x_i). \tag{6.1.17}$$

据第五章 (5.3.4) 式, 与边值问题 (6.1.1) 相应的双线性形式为

$$a(u, v) = \int_a^b [pu'v' + quv]\mathrm{d}x. \tag{6.1.18}$$

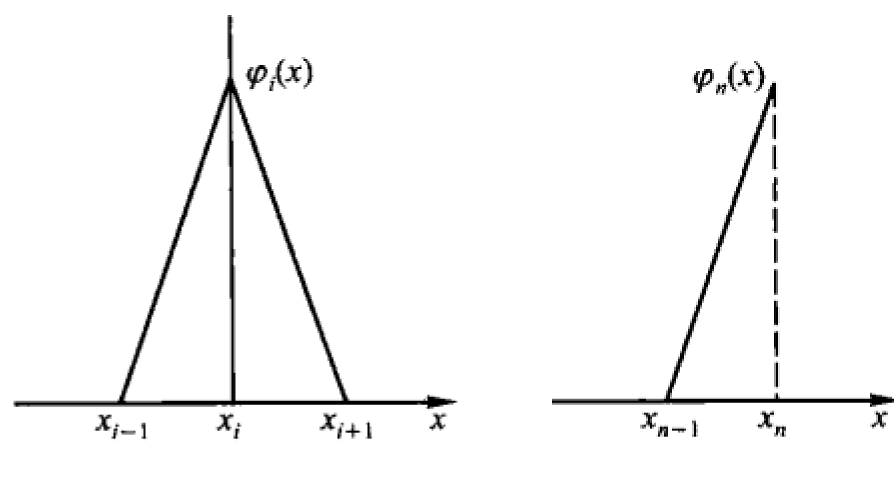


图 1.2

从而 Galerkin 方程为

$$\sum_{i=1}^{n} a(\varphi_i, \varphi_j) u_i = \int_a^b f \varphi_j \mathrm{d}x, j = 1, 2, \cdots, n. \tag{6.1.19}$$

若借助仿射变换 (6.1.3) 及 $[0,1]$ 上的标准山形函数

$$N_0(\xi) = 1 - \xi, \quad N_1(\xi) = \xi, \tag{6.1.20}$$

则对 $i = 1, 2, \cdots, n-1$, 基函数可写成:

$$\varphi_i(x) = \begin{cases} N_0(\xi),\ \xi = \dfrac{x - x_i}{h_{i+1}}, & x_i \leqslant x \leqslant x_{i+1}, \\ N_1(\xi),\ \xi = \dfrac{x - x_{i-1}}{h_i}, & x_{i-1} \leqslant x \leqslant x_i, \\ 0, & \text{其他}, \end{cases}$$

而

$$\varphi_n(x) = \begin{cases} N_1(\xi),\ \xi = \dfrac{x - x_{n-1}}{h_n}, & x_{n-1} \leqslant x \leqslant x_n, \\ 0, & \text{其他}. \end{cases}$$

显然, 当 $|j-i| \geqslant 2$ 时, $\varphi_i \cdot \varphi_j = 0$. 系数矩阵的第 j 行只有三个非零元素, 即

$$\begin{aligned} a(\varphi_{j-1}, \varphi_j) &= \int_{x_{j-1}}^{x_j} \left[p \varphi_j' \varphi_{j-1}' + q \varphi_j \varphi_{j-1}\right] \mathrm{d}x \\ &= \int_{x_{j-1}}^{x_j} \left[-p(x) h_j^{-2} + q(x) \varphi_j(x) \varphi_{j-1}(x)\right] \mathrm{d}x \\ &= \int_0^1 \left[-h_j^{-1} p(x_{j-1} + h_j \xi) + h_j q(x_{j-1} + h_j \xi)(1 - \xi)\xi\right] \mathrm{d}\xi, \end{aligned} \tag{6.1.21}$$

$$(6.1.22)\quad \begin{aligned} a(\varphi_j,\varphi_j) &= \int_{x_{j-1}}^{x_j} \left[p{\varphi'}_j^2 + q\varphi_j^2\right]\mathrm{d}x + \int_{x_j}^{x_{j+1}} \left[p{\varphi'}_j^2 + q\varphi_j^2\right]\mathrm{d}x \\ &= \int_{x_{j-1}}^{x_j} \left[p(x)h_j^{-2} + q(x)\varphi_j^2(x)\right]\mathrm{d}x + \\ &\quad \int_{x_j}^{x_{j+1}} \left[p(x)h_{j+1}^{-2} + q(x)\varphi_j^2(x)\right]\mathrm{d}x \\ &= \int_0^1 \left[h_j^{-1}p(x_{j-1}+h_j\xi) + h_j q(x_{j-1}+h_j\xi)\xi^2\right]\mathrm{d}\xi + \\ &\quad \int_0^1 \left[h_{j+1}^{-1}p(x_j+h_{j+1}\xi) + h_{j+1}q(x_j+h_{j+1}\xi)(1-\xi)^2\right]\mathrm{d}\xi, \end{aligned}$$

$$(6.1.23)\quad \begin{aligned} a(\varphi_{j+1},\varphi_j) &= \int_{x_j}^{x_{j+1}} \left[p\varphi_j'\varphi_{j+1}' + q\varphi_j\varphi_{j+1}\right]\mathrm{d}x \\ &= \int_{x_j}^{x_{j+1}} \left[-h_{j+1}^{-2}p(x) + q(x)\varphi_j(x)\varphi_{j+1}(x)\right]\mathrm{d}x \\ &= \int_0^1 \left[-h_{j+1}^{-1}p(x_j+h_{j+1}\xi) + h_{j+1}q(x_j+h_{j+1}\xi)\xi(1-\xi)\right]\mathrm{d}\xi. \end{aligned}$$

这里 $j=2,3,\cdots,n-1$. 第一行只有两个非零元素: $a(\varphi_1,\varphi_1)$, $a(\varphi_1,\varphi_2)$. 第 n 行也只有两个非零元素: $a(\varphi_n,\varphi_{n-1})$ 和

$$(6.1.24)\quad a(\varphi_n,\varphi_n) = \int_0^1 [h_n^{-1}p(x_{n-1}+h_n\xi) + h_n q(x_{n-1}+h_n\xi)\xi^2]\mathrm{d}\xi.$$

方程 (6.1.19) 的右端项

$$(6.1.25)\quad \begin{aligned} (f,\varphi_j) &= \int_{x_{j-1}}^{x_j} f(x)\varphi_j(x)\mathrm{d}x + \int_{x_j}^{x_{j+1}} f(x)\varphi_j(x)\mathrm{d}x \\ &= h_j\int_0^1 f(x_{j-1}+h_j\xi)\xi\mathrm{d}\xi + h_{j+1}\int_0^1 f(x_j+h_{j+1}\xi)(1-\xi)\mathrm{d}\xi. \end{aligned}$$

与 (6.1.9), (6.1.11) 比较, 可见 (6.1.19) 的系数矩阵就是总刚度矩阵.

由上看出, 按 Galerkin 方法推导有限元方程更加方便直接. 尤其重要的是, 因为它基于虚功原理, 所以不但可用于保守场问题, 也可用于非保守场及非驻定问题. 今后我们主要按这一观点建立有限元方程.

注 1.3 假若左边值条件非齐次

$$u(a) = \alpha,$$

则增加一基函数

$$\varphi_0(x) = \begin{cases} 1 - \dfrac{x-x_0}{h_1}, & x_0 \leqslant x \leqslant x_1, \\ 0 & \text{其他}. \end{cases}$$

将 u_h 表为 (参看第五章 §4 注 4.1)

$$u_h = \alpha\varphi_0(x) + \sum_{i=1}^{n} u_i\varphi_i(x),$$

自由度还是 n, 有限元方程为

$$\alpha a(\varphi_0, \varphi_j) + \sum_{i=1}^{n} a(\varphi_i, \varphi_j)u_i = (f, \varphi_j), \quad j = 1, 2, \cdots, n$$

(参看第五章 (5.4.14)), 或

$$\sum_{i=1}^{n} a(\varphi_i, \varphi_j)u_i = (f, \varphi_j) - \alpha a(\varphi_0, \varphi_j), \quad j = 1, 2, \cdots, n. \tag{6.1.26}$$

实际上只修改了第一个方程的右端, 因为当 $j = 2, 3, \cdots, n$ 时, $a(\varphi_0, \varphi_j) = 0$.

注 1.4 假若右边值条件非齐次

$$u'(b) = \beta,$$

则右端修改为 $(f, \varphi_j) + p(b)\beta\varphi_j(b)$ (参看第五章 §5 注 5.1), 有限元方程变成

$$\sum_{i=1}^{n} u_i \int_a^b [p\varphi_i'\varphi_j' + q\varphi_i\varphi_j]\mathrm{d}x = (f, \varphi_j) + p(b)\beta\varphi_j(b), \quad j = 1, 2, \cdots, n,$$

也只是修改了右端项, 实际上只修改了第 n 个方程的右端项, 因为当 $j = 1, 2, \cdots, n-1$ 时, $\varphi_j(b) = 0$.

1.3 收敛性和误差估计

现在给出有限元解的误差估计. 这里关键的一步是第五章定理 5.1 建立的不等式

$$||u - u_n||_1 \leqslant \beta \inf_{v \in U_n} ||u - v||_1.$$

应用到线性元就是

$$||u - u_h||_1 \leqslant \beta \inf_{v_h \in U_h} ||u - v_h||_1, \tag{6.1.27}$$

其中 u 是两点边值问题的解,u_h 是有限元解,U_h 是分段线性连续函数空间, $h = \max_i h_i$. 不等式 (6.1.27) 的几何意义是: 在空间 H^1 的度量下, 精确解和有限元解的距离不超过它和子空间 U_h 的最小距离乘以与 h 无关的常数因子, 而 $u_h \in U_h$, 所以估计式 (6.1.27) 是最佳的. 有了这个不等式, 就把估计 $||u - u_h||_1$ 化成用空间 U_h 逼近 u 这样一个纯属逼近论的问题了.

一种自然的逼近方法是, 在 U_h 中取 u 的插值函数 u_I (现在是分段线性插值) 作为逼近 u 的函数. 此时

$$\|u-u_h\|_1 \leqslant \beta\|u-u_I\|_1. \tag{6.1.28}$$

这归结为估计 u 的插值逼近误差. 因此建立 H^1 空间的插值理论是建立有限元理论的第二个关键. 对于我们现在的问题, 这一步是简单的.

假定解 u 于 $[a,b]$ 有连续二阶导数. 在任一单元 $I_i=[x_{i-1},x_i]$ 内考虑差

$$e(x)=u(x)-u_I(x).$$

显然 $e(x_{i-1})=e(x_i)=0$. 由 Rolle 定理, 有 $\xi\in I_i$, 使 $e'(\xi)=0$, 于是

$$\begin{aligned} e'(x) &= \int_\xi^x e''(t)\mathrm{d}t=\int_\xi^x u''(t)\mathrm{d}t, x\in I_i,\\ |e'(x)| &\leqslant \int_{x_{i-1}}^{x_i}|u''(t)|\mathrm{d}t\leqslant h_i^{1/2}\left(\int_{x_{i-1}}^{x_i}|u''(t)|^2\right)^{1/2} \end{aligned}$$

从而

$$\int_{x_{i-1}}^{x_i}|e'(t)|^2\mathrm{d}t\leqslant h_i^2\int_{x_{i-1}}^{x_i}|u''(t)|^2\mathrm{d}t,$$

这样便得插值误差估计

$$\|u-u_I\|_1\leqslant Ch\|u''\|,$$

右端 $\|\cdot\|$ 是 L^2- 模. 这里我们用了 Poincaré不等式. 与 (6.1.28) 结合即得有限元解的估计:

$$\|u-u_h\|_1\leqslant Ch\|u''\|,$$

这里以及以后常用 C 表示与 h,u 无关的常数. 若注意到对 $x\in I_i$,

$$\begin{aligned} (u-u_h)^2 &= \left(\int_\xi^x (u-u_h)'\mathrm{d}t\right)^2\leqslant (x-\xi)\int_{x_{i-1}}^{x_i}|(u-u_h)'|^2\mathrm{d}t\\ &\leqslant h\|u-u_h\|_1^2\leqslant Ch^3\|u''\|^2, \end{aligned}$$

则还可得到

$$\max_{a\leqslant x\leqslant b}|u-u_h|\leqslant Ch^{3/2}\|u''\|.$$

可见当 $h\to 0$ 时 u_h 一致趋于 u.

习 题

1. (实习题) 用线性元求下列边值问题的数值解:

$$-y'' + \frac{\pi^2}{4}y = \frac{\pi^2}{2}\sin\frac{\pi}{2}x, \quad 0 < x < 1, y(0) = 0, \quad y'(1) = 0$$

(精确解 $y = \sin\frac{\pi}{2}x$)

2. 就非齐第三边值条件 $u'(a) + \alpha_1 u(a) = \beta_1, u'(b) + \alpha_2 u(b) = \beta_2$ 导出 $-\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}u}{\mathrm{d}x}\right) + qu = f$ 的有限元方程.

§2 一维高次元

为了提高有限元法的精度, 需要增加试探函数空间 U_h 的维数. 这有两个途径, 一是加密网格剖分使单元最大直径 h 变小, 节点参数 $\{u_i\}$ 增加. 二是增加分段多项式的次数, 这就是本节要介绍的高次元. 引进高次元是有限元法的重要技巧.

一次元是分段一次多项式, 在每一单元 $I_i = [x_{i-1}, x_i]$ 上含有两个待定系数, 自由度是 2, 恰好由两个端点值决定. 分段二次、三次及高次多项式在每一单元上的自由度增加了, 应当按哪种插值去确定它们呢? 一种是 Lagrange 型, 在单元内部增加插值节点; 另一种是 Hermite 型, 在节点引进高阶导数. 无论用哪一种插值, 都要求它们在整个区间上有一定的光滑度, 以保证双线性形式有意义.

现在来构造试探函数空间. 像 §1 那样, 通过仿射变换

$$\xi = \frac{x - x_i}{h_{i+1}} \tag{6.2.1}$$

将节点 x_i 的右侧单元 $I_{i+1} = [x_i, x_{i+1}]$ 变到参考区间 $[0,1], x_i \to 0, x_{i+1} \to 1$. 而仿射变换

$$\xi = \frac{x_i - x}{h_i} \tag{6.2.2}$$

则把 x_i 的左侧单元 $I_i = [x_{i-1}, x_i]$ 变到 $[0,1], x_i \to 0, x_{i-1} \to 1$. 若把 (6.2.1) 改写成

$$\xi = \frac{|x - x_i|}{h_{i+1}}, \tag{6.2.3}$$

则只要用 I_i 的长度 h_i 替换 I_{i+1} 的长度 h_{i+1}, 就得到 (6.2.2). 作变换的目的, 是为了把任一单元上的插值归结到 $[0,1]$ 上的插值.

2.1 一次元 (线性元)

试探函数空间 U_h 由一切连续的折线函数组成 (满足本质边界条件), 就是说, U_h 中的函数在每一单元上线性, 在相邻单元的公共节点上连续. 一次函数有两个待定系数, 自由度是 2, 插值条件是: 在单元的两个端点取指定值.

为了构造节点基函数, 先在 $[0,1]$ 考虑插值问题: 求 $[0,1]$ 上的一次函数 $N_0(\xi)$, 使

$$N_0(0)=1, \quad N_0(1)=0.$$

显然, $N_0(\xi)=1-\xi$. 利用变换 (6.2.3), 便得到基函数 $\varphi_i(x)$ 在 $[x_i,x_{i+1}]$ 上的表达式

$$\varphi_i(x)=1-h_{i+1}^{-1}|x-x_i|.$$

若将 h_{i+1} 换成 h_i, 又得到 $\varphi_i(x)$ 在 $[x_{i-1},x_i]$ 上的表达式

$$\varphi_i(x)=1-h_i^{-1}|x-x_i|.$$

总之

$$\varphi_i(x)=\begin{cases}1-h_i^{-1}|x-x_i|, & x_{i-1}\leqslant x\leqslant x_i,\\ 1-h_{i+1}^{-1}|x-x_i|, & x_i\leqslant x\leqslant x_{i+1},\\ 0, & \text{其他}.\end{cases} \tag{6.2.4}$$

$\varphi_i(x)$ 是和节点 x_i 相应的**基函数**, 它们组成 U_h 的基底. 显然, $\varphi_i(x)$ 于节点连续. 若注意到 $\varphi_i(x)$ 在每个单元属于 C^1, 则知 $U_h\subset H^1(I)$.

在 §1 我们已详细讨论过一次有限元的构造及应用.

2.2 二次元

在每一单元上是二次多项式, 在单元节点处连续. 二次多项式有三个待定系数, 自由度是 3, 应给三个插值条件, 其中两个条件是在端点处取指定值, 另一个条件为在单元中点取指定值. 这样, 在每一单元中点增设了一个自由度, 总自由度几乎增加一倍.

基函数分两类, 一类对应节点, 另一类对应单元中点. 先在区间 $[0,1]$ 构造二次多项式 $N_0(\xi)$, 满足插值条件:

$$N_0(0)=1, \quad N_0\left(\frac{1}{2}\right)=N_0(1)=0.$$

显然, $N_0(\xi)$ 形如

$$N_0(\xi)=c\left(\xi-\frac{1}{2}\right)(\xi-1),$$

由 $N_0(0)=\frac{1}{2}c=1$, 知 $c=2$, 所以

(6.2.5) $$N_0(\xi)=(2\xi-1)(\xi-1).$$

用 (6.2.3) 消去 ξ, 即得

$$\varphi_i(x)=(2h_{i+1}^{-1}|x-x_i|-1)(h_{i+1}^{-1}|x-x_i|-1),\quad x_i\leqslant x\leqslant x_{i+1}.$$

这是 $\varphi_i(x)$ 的右半支. 若用 h_i 替换 h_{i+1}, 则得它的左半支. 总之

(6.2.6) $$\varphi_i(x)=\begin{cases}(2h_i^{-1}|x-x_i|-1)(h_i^{-1}|x-x_i|-1), & x_{i-1}\leqslant x\leqslant x_i,\\ (2h_{i+1}^{-1}|x-x_i|-1)(h_{i+1}^{-1}|x-x_i|-1), & x_i\leqslant x\leqslant x_{i+1},\\ 0, & \text{其他}.\end{cases}$$

其次, 构造二次多项式 $N_{\frac{1}{2}}(\xi)$, 它满足插值条件:

$$N_{\frac{1}{2}}(0)=N_{\frac{1}{2}}(1)=0,\quad N_{\frac{1}{2}}\left(\frac{1}{2}\right)=1.$$

显然

(6.2.7) $$N_{\frac{1}{2}}(\xi)=4\xi(1-\xi).$$

利用 (6.2.1) 消去 ξ, 就给出和 $x_{i+\frac{1}{2}}=x_i+\frac{1}{2}h_{i+1}$ 相关的基函数

(6.2.8) $$\varphi_{i+\frac{1}{2}}(x)=\begin{cases}4h_{i+1}^{-1}(x-x_i)(1-h_{i+1}^{-1}(x-x_i)), & x_i\leqslant x\leqslant x_{i+1},\\ 0, & \text{其他}.\end{cases}$$

以所有 $\varphi_i,\varphi_{i+\frac{1}{2}}$ 为基底, 张成二次元试探函数空间 U_h. 任一 $u_h\in U_h$ 可唯一表成

$$u_h=\sum_i[u_i\varphi_i(x)+u_{i+\frac{1}{2}}\varphi_{i+\frac{1}{2}}(x)].$$

这是 Lagrange 型插值, 试探函数 $u_h(x)$ 在节点 x_i 处连续. 若注意到 $u_h(x)$ 在每一单元上属于 C^1, 则知 $U_h\subset H^1(I)$.

在 x_i 和 $x_{i+\frac{1}{2}}$ 处的有限元方程是不同的. 节点 x_i 处的方程联系五个未知量: $u_{i-1},u_{i-\frac{1}{2}},u_i,u_{i+\frac{1}{2}},u_{i+1}$; 而在半整数点 $x_{i+\frac{1}{2}}$ 的方程只联系三个未知量 $u_i,u_{i+\frac{1}{2}}$, u_{i+1}. 因此, 解有限元方程时可先用半整数点的方程消去 $u_{i+\frac{1}{2}}$, 再解一个阶数几乎小一半的方程组.

图 2.1 就单元 $[-h,0],[0,h]$ 画出了 Lagrange 二次元的基函数 $\varphi_0(x),\varphi_{\frac{1}{2}}(x)$ 的几何图形.

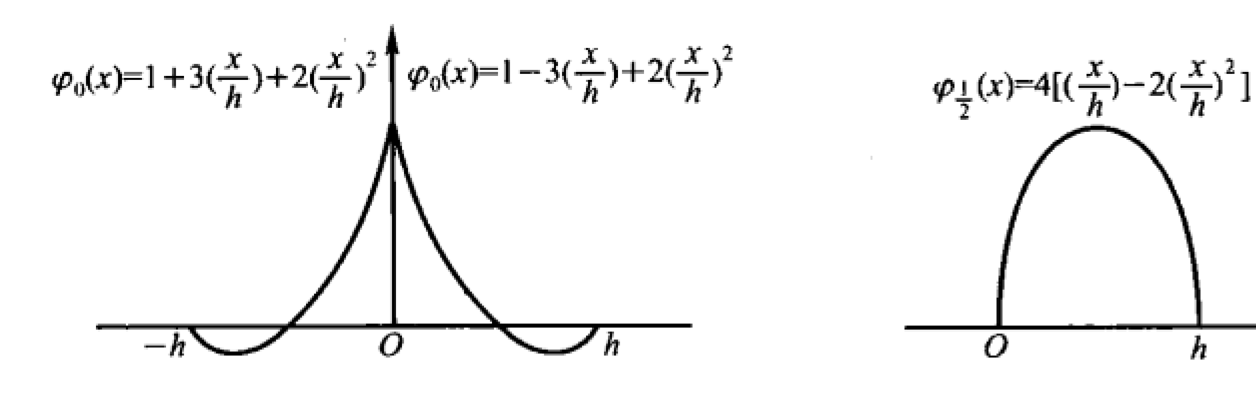


图 2.1 Lagrange 二次元基函数曲线图

2.3 三次元

要想得到 $C^1(I)$ 类试探函数空间, 可采用分段三次多项式函数. 三次多项式有四个待定系数, 自由度等于 4. 插值条件是: 在两个端点取指定的函数值和一阶导数. 这是 Hermite 型插值, 每一节点对应两个基函数.

首先在 $[0,1]$ 求三次多项式 $N_0(\xi)$, 使之满足条件:

$$N_0(0)=1,\quad N_0'(0)=0,\quad N_0(1)=N_0'(1)=0.$$

显然, $N_0(\xi)$ 形如

$$N_0(\xi)=(1-\xi)^2(\alpha\xi+\beta).$$

为使前两个条件满足, 应取 $\alpha=2,\beta=1$. 于是

$$N_0(\xi)=(1-\xi)^2(2\xi+1). \tag{6.2.9}$$

利用变换 (6.2.3) 消去 ξ, 就得到第一个基函数

$$\varphi_i^{(0)}(x)=\begin{cases}(1-h_i^{-1}|x-x_i|)^2(2h_i^{-1}|x-x_i|+1), & x_{i-1}\leqslant x\leqslant x_i,\\ (1-h_{i+1}^{-1}|x-x_i|)^2(2h_{i+1}^{-1}|x-x_i|+1), & x_i\leqslant x\leqslant x_{i+1},\\ 0, & \text{其他}.\end{cases} \tag{6.2.10}$$

其次, 求三次多项式 $N_1(\xi)$, 满足条件

$$N_1(0)=0,\quad \left.\frac{\mathrm{d}N_1(\xi)}{\mathrm{d}x}\right|_{\xi=0}=1,\quad N_1(1)=N_1'(1)=0.$$

显然, $N_1(\xi)$ 形如

$$N_1(\xi)=c\xi(1-\xi)^2.$$

常数 c 应如此选取, 使

$$\left.\frac{\mathrm{d}N_1}{\mathrm{d}x}\right|_{x=x_i}=\frac{\mathrm{d}N_1}{\mathrm{d}\xi}\cdot\left.\frac{\mathrm{d}\xi}{\mathrm{d}x}\right|_{x=x_i}=\begin{cases}h_{i+1}^{-1}\left.\dfrac{\mathrm{d}N_1}{\mathrm{d}\xi}\right|_{\xi=0}=ch_{i+1}^{-1}, & x_i\leqslant x\leqslant x_{i+1},\\ -h_i^{-1}\left.\dfrac{\mathrm{d}N_1}{\mathrm{d}\xi}\right|_{\xi=0}=-ch_i^{-1}, & x_{i-1}\leqslant x\leqslant x_i\end{cases}=1,$$

从而

$$c=\begin{cases}h_{i+1}, & x_i\leqslant x\leqslant x_{i+1},\\ -h_i, & x_{i-1}\leqslant x\leqslant x_i.\end{cases}$$

利用 (6.2.1), (6.2.2) 及 (6.2.3) 消去 ξ, 得到第二个基函数:

$$\varphi_i^{(1)}(x)=\begin{cases}(x-x_i)(h_i^{-1}|x-x_i|-1)^2, & x_{i-1}\leqslant x\leqslant x_i,\\ (x-x_i)(h_{i+1}^{-1}|x-x_i|-1)^2, & x_i\leqslant x\leqslant x_{i+1},\\ 0, & \text{其他}.\end{cases}\tag{6.2.11}$$

以所有 $\varphi_i^{(0)}(x),\varphi_i^{(1)}(x)$ 为基底, 张成一个三次元试探函数空间 U_h. 显然, 任一 $u_h\in U_h$ 属于 $C^1(I)$. 若注意到 u_h 分段二次连续可微, 则知 $U_h\subset H^2(I)$. 任一 u_h 可表成

$$u_h=\sum_i[u_i\varphi_i^{(0)}(x)+u_i'\varphi_i^{(1)}(x)],\tag{6.2.12}$$

其中 $u_i=u_h(x_i),u_i'=u_h'(x_i)$.

第二个基函数也可以不必通过仿射变换, 而直接用待定系数法确定. 因 $\varphi_i^{(1)}(x_i)$ 在 $[x_i,x_{i+1}]$ 满足插值条件

$$\varphi_i^{(1)}(x_i)=0,\qquad \left.\frac{\mathrm{d}}{\mathrm{d}x}\varphi_i^{(1)}(x)\right|_{x=x_i}=1,$$
$$\varphi_i^{(1)}(x_{i+1})=0,\qquad \left.\frac{\mathrm{d}}{\mathrm{d}x}\varphi_i^{(1)}(x)\right|_{x=x_{i+1}}=0.$$

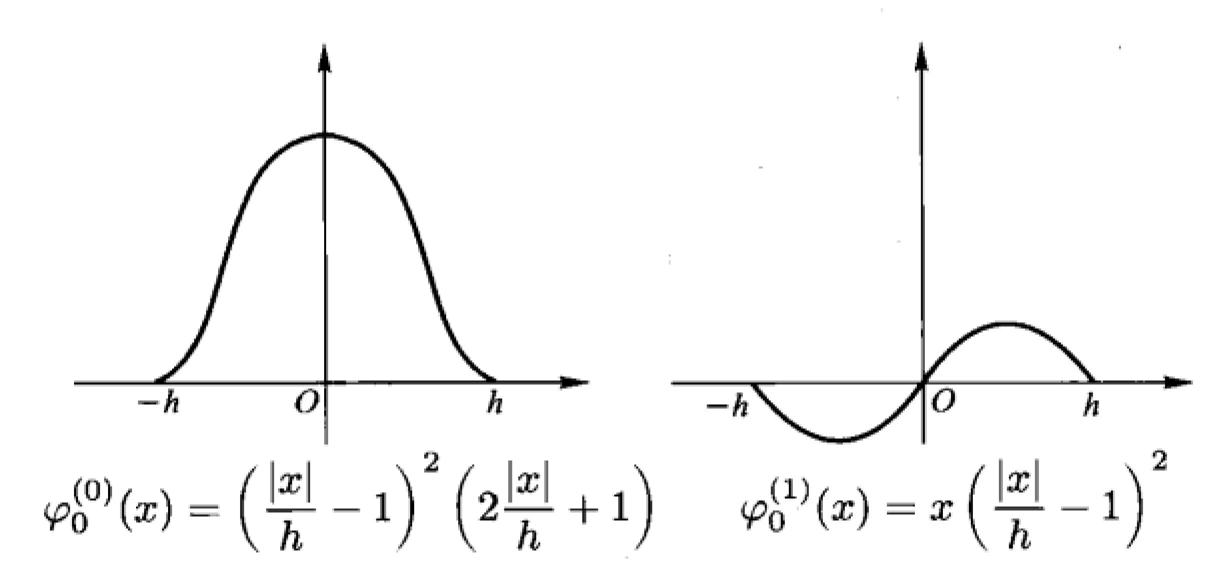
故可直接写出

$$\varphi_i^{(1)}(x)=c(x-x_i)(x-x_{i+1})^2.$$

据第二个插值条件, 定出 $c=h_{i+1}^{-2}$. 依次用 h_i,x_{i-1} 替代 h_{i+1},x_{i+1}, 即得 $\varphi_i^{(1)}(x)$ 的左半支. 总之

$$\varphi_i^{(1)}(x)=\begin{cases}h_i^{-2}(x-x_i)(x-x_{i-1})^2, & x_{i-1}\leqslant x\leqslant x_i,\\ h_{i+1}^{-2}(x-x_i)(x-x_{i+1})^2, & x_i\leqslant x\leqslant x_{i+1},\\ 0, & \text{其他}.\end{cases}$$

图 2.2 就单元 $[-h,0],[0,h]$ 画出了 Hermite 三次元的基函数 $\varphi_0^{(0)}(x),\varphi_0^{(1)}(x)$ 的几何图形.



$$\varphi_0^{(0)}(x)=\left(\frac{|x|}{h}-1\right)^2\left(2\frac{|x|}{h}+1\right)\qquad \varphi_0^{(1)}(x)=x\left(\frac{|x|}{h}-1\right)^2$$

图 2.2 Hermite 三次元基函数曲线图

以上方法, 同样可用来构造四次元、五次元以及更高次元.

做出基函数后, 还要形成有限元方程, 这和线性有限元基本相同, 只是工作量更大些罢了.

例如用三次有限元解两点边值问题 (6.1.1), 有限元方程为

$$a(u_h,\varphi_j^{(l)})=(f,\varphi_j^{(l)}),\quad l=0,1,\quad j=1,2,\cdots,n.$$

将 u_h 的表达式 (6.2.12) 代入, 得

$$\sum_{i=0}^{n}\left[u_i a(\varphi_i^{(0)},\varphi_j^{(l)})+u_i' a(\varphi_i^{(1)},\varphi_j^{(l)})\right]=(f,\varphi_j^{(l)}),\tag{6.2.13}$$

$$l=0,1,\quad j=1,2,\cdots,n,$$

左端点方程为

$$\begin{cases}u_0=0,\\ \displaystyle\sum_{i=0}^{n}\left[u_i a(\varphi_i^{(0)},\varphi_0^{(1)})+u_i' a(\varphi_i^{(1)},\varphi_0^{(1)})\right]=(f,\varphi_0^{(1)}).\end{cases}\tag{6.2.14}$$

为形成系数矩阵, 要计算大量较复杂的积分. 一般要用数值积分, 通常用 Gauss 求积公式.

方程 (6.2.13) 有两组未知量:$\{u_i\}$ 和 $\{u_i'\}$ (称为**广义坐标**), 相当于固体力学中的位移和应力, 都是实际要计算的量. 假若用一次元, 求出 u_i 后, 还要作一次微商运算 (左、右微商) 和加权平均, 才得到 u_i'. 这样, 往往会影响应力的精度. 用高次元则可改善应力的计算.

高次元的另一个优点是收敛阶高. 例如三次元, 若精确解足够光滑, 则收敛阶可达到 $O(h^3)$ (按 H^1 度量), 这就可以适当放大步长.

另一方面也要看到, 采用高次元要付出一些代价. 除了增加计算积分的复杂性外, 刚度矩阵的带宽也比一次元更大了.

此外, 若边值问题的解本身不够光滑, 用高次元就不能达到高精度的目的. 所以, 用哪种类型的试探函数作有限元逼近, 要根据问题的性质和机器条件决定.

习　　题

1. 证明下列分段二次多项式基函数属于 $C^1([0,1])$, 并作出其图形.

$$\varphi_i^{(0)}(x)=\begin{cases}0, & 0\leqslant x\leqslant x_{i-1}, x_{i+1}\leqslant x\leqslant 1,\\ 2\left(\dfrac{x-x_{i-1}}{h}\right)^2, & x_{i-1}\leqslant x\leqslant x_i-0.5h,\\ -2\left(\dfrac{x-x_i}{h}\right)^2+1, & x_i-0.5h\leqslant x\leqslant x_i+0.5h,\\ 2\left(\dfrac{x-x_{i+1}}{h}\right)^2, & x_i+0.5h\leqslant x\leqslant x_{i+1};\end{cases}$$

$$\varphi_i^{(1)}(x)=\begin{cases}0, 0\leqslant x\leqslant x_{i-1}, & x_{i+1}\leqslant x\leqslant 1,\\ -\dfrac{1}{2h}(x-x_{i-1})^2, & x_{i-1}\leqslant x\leqslant x_i-0.5h,\\ x-x_i+\dfrac{3}{2h}(x-x_i)^2, & x_i-0.5h\leqslant x\leqslant x_i,\\ x-x_i-\dfrac{3}{2h}(x-x_i)^2, & x_i\leqslant x\leqslant x_i+0.5h,\\ \dfrac{1}{2h}(x-x_{i+1})^2, & x_i+0.5h\leqslant x\leqslant x_{i+1}.\end{cases}$$

2. 试就 $p=1, q=0, f=1$ 具体写出有限元方程 (6.2.13),(6.2.14).

§3　解二维问题的矩形元

从本节起, 我们讨论二维椭圆边值问题的有限元解法. 首先讨论一些常用单元的形状函数及构造方法, 包括矩形元、三角形元和曲边形元. 本节讨论矩形元.

3.1　Lagrange 型公式

为简单起见, 假定区域 G 可以分割成有限个矩形的和, 且每个小矩形的边和坐标轴平行. 任意两个矩形, 或者不相交, 或者有公共的边或公共的顶点. 我们把每一小矩形叫做**单元**, 称如此的分割为**矩形剖分**.

取定剖分后, 我们着手构造相应的试探函数 $u_h(x,y)$ (h 为单元的最大直径), 它在每一单元上是多项式 —— 称为**单元形状函数**, 在相邻单元之间有一定的光滑性. 如果造出了单元形状函数, 试探函数也就得到了.

任一个矩形 $R_{ij}=[x_i,x_j+\Delta x;y_j,y_j+\Delta y]=\{x_i\leqslant x\leqslant x_j+\Delta x;y_j\leqslant y\leqslant y_j+\Delta y\}$, 通过仿射变换

$$\xi=(x-x_i)/\Delta x,\quad \eta=(y-y_j)/\Delta y, \tag{6.3.1}$$

总可以变成正方形 $I\times I=[0,1]\times[0,1]$. 只要在 $I\times I$ 上造出了形状函数, 再通过变换 (6.3.1) 就得到 R_{ij} 上的形状函数.

$I\times I$ 上最简单的形状函数是双线性函数

$$P_{11}(\xi,\eta)=C_0+C_1\xi+C_2\eta+C_3\xi\eta, \tag{6.3.2}$$

其系数由 u 在四个顶点上的值 $u_{00},u_{01},u_{10},u_{11}$ 决定 (插值节点分布如图 3.1).

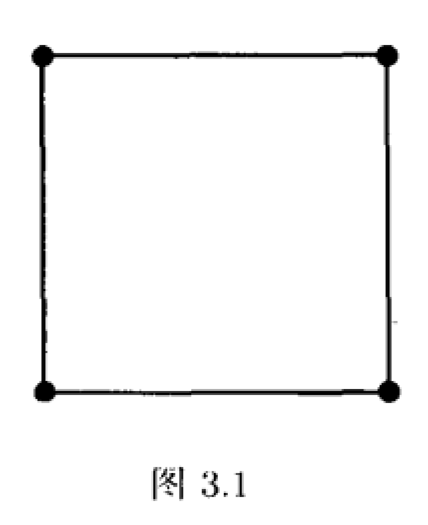


图 3.1

利用一维山形函数

$$N_0(\xi)=1-\xi,\quad N_1(\xi)=\xi, \tag{6.3.3}$$

可将 $P_{11}(\xi,\eta)$ 表示成

$$\begin{aligned}P_{11}(\xi,\eta)=&N_0(\xi)N_0(\eta)u_{00}+N_0(\xi)N_1(\eta)u_{01}+\\&N_1(\xi)N_0(\eta)u_{10}+N_1(\xi)N_1(\eta)u_{11}.\end{aligned} \tag{6.3.4}$$

因右端各项插值基函数是一元插值基函数的乘积, 故称 (6.3.4) 为**乘积型二元插值公式**. 对于乘积型区域, 都可以由一维插值公式组合成高维乘积型插值公式, 这一技巧, 后边还要多次用到.

通过变换 (6.3.1) 消去 ξ,η, 便得到 R_{ij} 上的形状函数. 把这些函数拼凑在一起, 就是 G 上的试探函数. 实际计算时, 并不需要消去中间变量 ξ,η, 因为计算刚度矩阵元素时 (定积分) 用 ξ,η 作自变量更加方便.

显然, 试探函数由节点上的一组 $\{u_{ij}\}$ (称为**广义坐标**) 唯一决定, 因此试探函数空间 U_h 的维数和 $\{u_{ij}\}$ 的个数相等, 是有限维空间. 现在证明, 试探函数在相邻单元的公共边上连续. 因为在公共边上它是一个变量的线性函数, 且在两个

端点取相同值, 因此试探函数由两侧单元确定的边值一致. 这说明空间 $U_h \subset C$, 从而 $U_h \subset H^1$.

有了插值公式 (6.3.4), 基函数就可直接得到. 实际上, (6.3.4) 右端第一个基函数

$$N_0(\xi)N_0(\eta) = (1-\xi)(1-\eta) \tag{6.3.5}$$

在 $(0,0)$ 等于 1, 在其余三个顶点 $(0,1),(1,0),(1,1)$ 等于 0. 消去 ξ,η, 就得到和节点 (x_i,y_j) 相关的基函数在单元 R_{ij} 上的限制

$$\left(1-\frac{x-x_i}{\Delta x}\right)\left(1-\frac{y-y_j}{\Delta y}\right),$$

这里 $\Delta x,\Delta y>0$ 为 R_{ij} 的底和高的长度, $x>x_i, y>y_j$. 在其他三个以 (x_i,y_j) 为顶点的矩形上, 基函数仍保持如上形式, 只不过 $\Delta x,\Delta y$ 解释成相应的矩形的底和高的代数值, 例如在以 (x_i,y_j) 为原点的第二象限上, $\Delta x<0,\Delta y>0$, 同时 $x-x_i<0, y-y_j>0$. 若采用底和高的绝对长度 $|\Delta x|,|\Delta y|$, 则基函数也可表成

$$\varphi_{ij}(x,y)=\begin{cases}\left(1-\dfrac{|x-x_i|}{|\Delta x|}\right)\left(1-\dfrac{|y-y_j|}{|\Delta y|}\right), & (x,y)\in R_{ij},\\ 0 & \text{其他},\end{cases} \tag{6.3.6}$$

其中 R_{ij} 是以 (x_i,y_j) 为顶点的矩形单元.

为了得到 $I\times I$ 上的双二次函数:

$$P_{22}(\xi,\eta)=C_0+C_1\xi+C_2\eta+C_3\xi^2+C_4\xi\eta+C_5\eta^2+C_6\xi^2\eta+C_7\xi\eta^2+C_8\xi^2\eta^2,$$

应给出九个插值节点: $(0,0),\left(\frac{1}{2},0\right),(1,0),\left(0,\frac{1}{2}\right),\left(\frac{1}{2},\frac{1}{2}\right),\left(1,\frac{1}{2}\right),(0,1),\left(\frac{1}{2},1\right),(1,1)$ (参见图 3.2), 利用一维二次插值公式的基函数 (参看 §2)

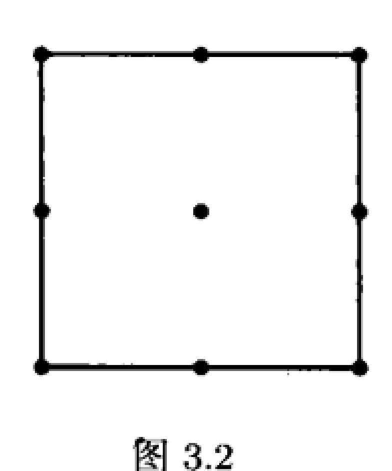


图 3.2

$$\begin{cases}N_0(\xi)=(2\xi-1)(\xi-1),\\ N_{\frac{1}{2}}(\xi)=4\xi(1-\xi),\\ N_1(\xi)=\xi(2\xi-1),\end{cases} \tag{6.3.7}$$

即得乘积型插值公式

$$P_{22}(\xi,\eta)=\sum_{i,j=0}^{2}N_{\frac{i}{2}}(\xi)N_{\frac{j}{2}}(\eta)u_{\frac{i}{2}\frac{j}{2}}. \tag{6.3.8}$$

通过变换 (6.3.1) 消去 ξ 和 η, 就得到 R_{ij} 上的形状函数.

运用前面的证法, 可知试探函数于 $\overline{G}$ 连续, 从而 $U_h\subset H^1$. 从 (6.3.8) 不难作出和节点相关的基函数.

同样可构造 $I\times I$ 上的双三次插值公式

$$P_{33}(\xi,\eta)=\sum_{i,j=0}^{3}N_{\frac{i}{3}}(\xi)N_{\frac{j}{3}}(\eta)u_{\frac{i}{3}\frac{j}{3}}, \tag{6.3.9}$$

其中

$$\begin{cases}N_0(\xi)=-\dfrac{9}{2}(\xi-1)\left(\xi-\dfrac{1}{3}\right)\left(\xi-\dfrac{2}{3}\right),\\ N_{\frac{1}{3}}(\xi)=\dfrac{27}{2}\xi(\xi-1)\left(\xi-\dfrac{2}{3}\right),\\ N_{\frac{2}{3}}(\xi)=-\dfrac{27}{2}\xi(\xi-1)\left(\xi-\dfrac{1}{3}\right),\\ N_1(\xi)=\dfrac{9}{2}\xi\left(\xi-\dfrac{1}{3}\right)\left(\xi-\dfrac{2}{3}\right).\end{cases} \tag{6.3.10}$$

插值节点分布如图 3.3. 相应的试探函数空间 $U_h\in H^1$.

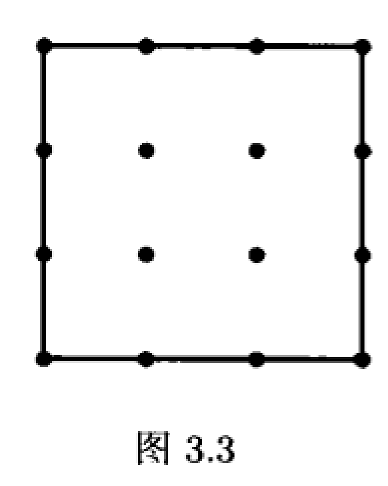


图 3.3

3.2　Hermite 型公式

前面得到的插值公式都是 Lagrange 型的, 广义坐标中不出现导数. 若采用 Hermite 型公式, 则广义坐标中出现导数.

最简单的 Hermite 型公式是: 取四个顶点为插值节点, 顶点上的 u,u_x,u_y 为广义坐标, 此时单元自由度等于 12. 而完全三次多项式 P_3 含 10 个系数. 为使插值问题有唯一解, 再加进 x^3y 和 xy^3 两项. 于是我们的插值问题为求形如

$$P_3(x,y)+\alpha x^3y+\beta xy^3 \tag{6.3.11}$$

的多项式, 在顶点取指定的 u, u_x 和 u_y. 因为 (6.3.11) 的形式在变换 (6.3.1) 下不变, 所以可把它转化成单位正方形 $I \times I$ 上的插值问题.

为此我们先构造基函数, 即求形如 (6.3.11) 的基函数 $N^i(\xi,\eta)(i=0,1,2)$, 满足插值条件:

(i) $N^{(0)}$ 在 $(0,0)$ 的值等于 1, 一阶偏导数等于 0, 在其他三个顶点的函数值及一阶偏导数值等于 0.

(ii) $N^{(1)}$ 在 $(0,0)$ 对 ξ 的一阶偏导数等于 Δx (相当于它对 x 的一阶偏导数等于 1, 参看变换 (6.3.1)), 对 η 的一阶偏导数等于 0, $N^{(1)}(0,0)=0$, 而在其余三个顶点的函数值和一阶偏导数等于 0.

(iii) $N^{(2)}$ 在 $(0,0)$ 对 η 的一阶偏导数等于 Δy (相当于它对 y 的一阶偏导数等于 1), 对 ξ 的一阶偏导数等于 0, $N^{(2)}(0,0)=0$, 而在其余三个顶点的函数值和一阶偏导数等于 0.

设

$$
\begin{aligned}
(6.3.12)\qquad N^{(0)}(\xi,\eta) = {} & C_0 + C_1\xi + C_2\eta + C_3\xi^2 + C_4\xi\eta + C_5\eta^2 + \\
& C_6\xi^3 + C_7\xi^2\eta + C_8\xi\eta^2 + C_9\eta^3 + C_{10}\xi^3\eta + C_{11}\xi\eta^3,
\end{aligned}
$$

则

$$
\begin{aligned}
N^{(0)}(\xi,0) &= C_0 + C_1\xi + C_3\xi^2 + C_6\xi^3, \\
N^{(0)}(0,\eta) &= C_0 + C_2\eta + C_5\eta^2 + C_9\eta^3.
\end{aligned}
$$

它们在 $[0,1]$ 上满足一维 Hermite 插值条件. 由 (6.2.9) 式, 它们分别等于 $(1-\xi)^2(2\xi+1)$ 和 $(1-\eta)^2(2\eta+1)$. 比较系数, 即知

$$
C_0 = 1, \quad C_1 = C_2 = 0, \quad C_3 = C_5 = -3, \quad C_6 = C_9 = 2.
$$

又由插值条件, 三次多项式

$$
\begin{aligned}
N^{(0)}(\xi,1) &= (C_4 + C_8 + C_{11})\xi + (-3 + C_7)\xi^2 + (2 + C_{10})\xi^3, \\
N^{(0)}(1,\eta) &= (C_4 + C_7 + C_{10})\eta + (-3 + C_8)\eta^2 + (2 + C_{11})\eta^3
\end{aligned}
$$

在 $[0,1]$ 上都恒等于 0, 因此

$$
C_4 = -1, \quad C_7 = C_8 = 3, \quad C_{10} = C_{11} = -2.
$$

总之,

$$
(6.3.13)\quad N^{(0)}(\xi,\eta) = 1 - 3(\xi^2+\eta^2) - \xi\eta + 2(\xi^3+\eta^3) + 3(\xi^2\eta + \xi\eta^2) - 2(\xi^3\eta + \xi\eta^3).
$$

用同样办法可得

$$N^{(1)}(\xi,\eta)=(\xi-2\xi^2-\xi\eta+\xi^3+2\xi^2\eta-\xi^3\eta)\Delta x. \tag{6.3.14}$$

$$N^{(2)}(\xi,\eta)=(\eta-2\eta^2-\xi\eta+\eta^3+2\xi\eta^2-\xi\eta^3)\Delta y. \tag{6.3.15}$$

再通过变换 (6.3.1), 就得到相应于节点 (x_i,y_j) 的三个基函数:

$$\varphi_{ij}^{(0)}=\begin{cases}N^{(0)}, & (x,y)\in R_{ij},\\ 0, & \text{其他}.\end{cases} \tag{6.3.16}$$

$$\varphi_{ij}^{(1)}=\begin{cases}N^{(1)}, & (x,y)\in R_{ij},\\ 0, & \text{其他}.\end{cases} \tag{6.3.17}$$

$$\varphi_{ij}^{(2)}=\begin{cases}N^{(2)}, & (x,y)\in R_{ij},\\ 0, & \text{其他}.\end{cases} \tag{6.3.18}$$

此处 R_{ij} 是以 (x_i,x_j) 为顶点的任一矩形 (共四个), Δx 和 Δy 是 R_{ij} 的底和高的代数值. 于是任一试探函数可唯一表为:

$$u(x,y)=\sum_{i,j}[u_{ij}\varphi_{ij}^{(0)}+(u_x)_{ij}\varphi_{ij}^{(1)}+(u_y)_{ij}\varphi_{ij}^{(2)}]. \tag{6.3.19}$$

容易验证, 试探函数属于 C, 且分片属于 C^1, 因此试探函数空间 $U_h\subset H^1$. 这一最简单形式的 Hermite 型元称为 Adini 元.

习　　题

1. 证明由双二次插值公式、双三次插值公式生成的试探函数空间属于 H^1.
2. 从插值公式 (6.3.8) 导出和节点相关的基函数.
3. 证明由 Adini 元确定的试探函数空间属于 H^1.

§4　三 角 形 元

对曲边区域 G, 一般采用三角网近似. 不妨设 G 是多边形域 (否则可用多边形域逼近它). 将 G 分割成有限个三角形之和, 使不同三角形无重叠的内部, 且任一三角形的顶点不属于其他三角形边的内部. 这样就把 G 分割成三角形网, 称为 G 的三角剖分. 每个三角形称为**单元**, 它的顶点称为**节点**. 属同一单元的二顶点称为**相邻节点**, 有公共边的两个三角形称为**相邻单元**.

由于三角剖分可构造非均匀网格, 并且能较好地逼近具复杂边界的区域, 所以在二维问题中, 三角形元是应用最广的单元.

4.1 面积坐标及有关公式

设 $\Delta(i,j,k)$ 是以 i,j,k 为顶点的任意三角形单元, 面积为 S. 我们约定 i,j,k 的次序按逆时针方向排列. 在 $\Delta(i,j,k)$ 内任取一点 P, 坐标为 (x,y). 过 P 点作与三个顶点的连线, 将 $\Delta(i,j,k)$ 分成三个三角形 (参见图 4.1): $\Delta(i,j,P),\Delta(j,k,P),\Delta(k,i,P)$, 其面积分别为 S_k,S_i,S_j. 显然 $S_i+S_j+S_k=S$. 令

$$L_i=S_i/S,\quad L_j=S_j/S,\quad L_k=S_k/S, \tag{6.4.1}$$

则 $L_i,L_j,L_k\geqslant 0, L_i+L_j+L_k=1$. 给定一点 P, 唯一确定如此的一组数 (L_i,L_j,L_k). 反之, 任给一组 $(L_i,L_j,L_k), L_i,L_j,L_k\geqslant 0, L_i+L_j+L_k=1$. 按关系式 (6.4.1) 也唯一确定一点 P. 所以同一点 P, 既可用直角坐标 (x,y) 表示, 也可用 (L_i,L_j,L_k) 表示. 我们称 (L_i,L_j,L_k) 为点 P 的**面积坐标**. 因为三角形的面积与参考坐标系无关, 所以面积坐标也与坐标系无关, 这是采用面积坐标的优点.

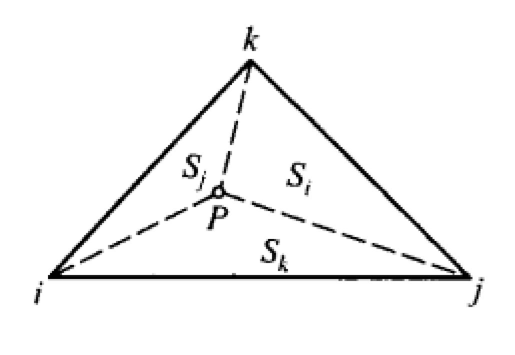


图 4.1

我们知道

$$\begin{cases} 2S=\begin{vmatrix}1 & x_i & y_i\\ 1 & x_j & y_j\\ 1 & x_k & y_k\end{vmatrix}, \quad 2S_i=\begin{vmatrix}1 & x & y\\ 1 & x_j & y_j\\ 1 & x_k & y_k\end{vmatrix} \\ 2S_j=\begin{vmatrix}1 & x_i & y_i\\ 1 & x & y\\ 1 & x_k & y_k\end{vmatrix}, \quad 2S_k=\begin{vmatrix}1 & x_i & y_i\\ 1 & x_j & y_j\\ 1 & x & y\end{vmatrix}. \end{cases} \tag{6.4.2}$$

由此可建立面积坐标与直角坐标之间的下列转换关系:

$$\begin{cases} L_i=\dfrac{1}{2S}[(x_jy_k-x_ky_j)+(y_j-y_k)x+(x_k-x_j)y], \\ L_j=\dfrac{1}{2S}[(x_ky_i-x_iy_k)+(y_k-y_i)x+(x_i-x_k)y], \\ L_k=\dfrac{1}{2S}[(x_iy_j-x_jy_i)+(y_i-y_j)x+(x_j-x_i)y]. \end{cases} \tag{6.4.3}$$

$$
\begin{cases} x = x_iL_i + x_jL_j + x_kL_k, \\ y = y_iL_i + y_jL_j + y_kL_k. \end{cases} \tag{6.4.4}
$$

在推导后一关系式时, 利用了等式 $L_i + L_j + L_k = 1$.

由连锁规则不难看出

$$
\begin{cases} \dfrac{\partial}{\partial x} = \dfrac{1}{2S}\left[(y_j - y_k)\dfrac{\partial}{\partial L_i} + (y_k - y_i)\dfrac{\partial}{\partial L_j} + (y_i - y_j)\dfrac{\partial}{\partial L_k}\right], \\ \dfrac{\partial}{\partial y} = \dfrac{1}{2S}\left[(x_k - x_j)\dfrac{\partial}{\partial L_i} + (x_i - x_k)\dfrac{\partial}{\partial L_j} + (x_j - x_i)\dfrac{\partial}{\partial L_k}\right]. \end{cases} \tag{6.4.5}
$$

面积坐标 L_i, L_j, L_k 满足

$$
L_i + L_j + L_k = 1.
$$

从中解出 $L_i = 1 - L_j - L_k$, 消去 (6.4.4) 右端的 L_i, 则得到由 L_jL_k 平面到 xy 平面的仿射变换, 其逆变换把 $\Delta(i,j,k)$ 变到 L_jL_k 平面以 $(0,0),(0,1),(1,0)$ 为顶点的直角三角形, 如图 4.2,

$$
\begin{aligned} (x_i, y_i) &\leftrightarrow (0,0), \\ (x_j, y_j) &\leftrightarrow (1,0), \\ (x_k, y_k) &\leftrightarrow (0,1). \end{aligned}
$$

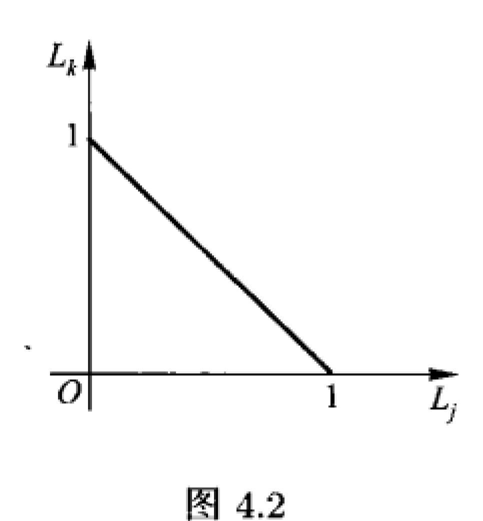


图 4.2

利用重积分变量替换公式, 不难得出下列积分公式

$$
\iint\limits_{\Delta(i,j,k)} L_i^p L_j^q L_k^r \mathrm{d}x\mathrm{d}y = 2S\frac{p!q!r!}{p+q+r+2}, \tag{6.4.6}
$$

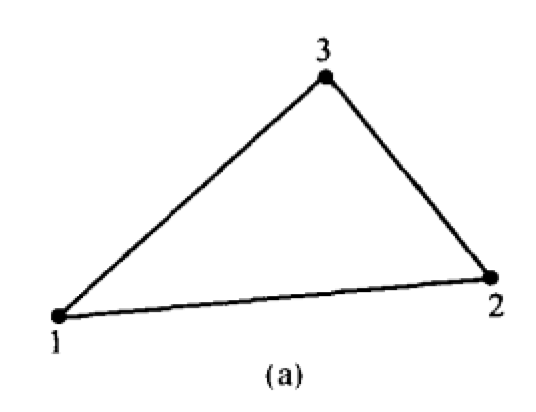
其中 p, q, r 是任意非负整数.

4.2 Lagrange 型公式

在三角形元 $\Delta(1,2,3)$ 上, 要构造一个 m 次完全多项式

$$p_m(x,y)=\sum_{i+j=0}^{m}c_{ij}x^iy^j, \tag{6.4.7}$$

用它去插值逼近 $u(x,y)$. 因为此多项式含有 $\frac{1}{2}(m+1)(m+2)$ 个待定系数, 所以在 $\Delta(1,2,3)$ 上应取同样个数的插值节点.



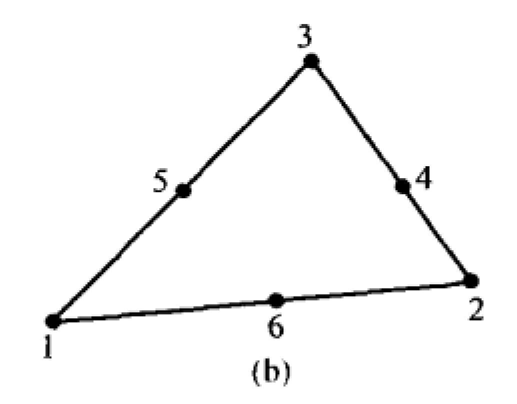


图 4.3

$m=1$ 时, $p_1(x,y)$ 是一次多项式, 插值节点数是 3. 取 $\Delta(1,2,3)$ 的三个顶点为插值节点 (参看图 4.3(a)), 用待定系数法, 易得

$$P_1(x,y)=L_1u_1+L_2u_2+L_3u_3, \tag{6.4.8}$$

此处 (L_1,L_2,L_3) 是 (x,y) 的面积坐标.

$m=2$ 时, $P_2(x,y)$ 是二次多项式, 插值节点应取 6 个, 通常取 $\Delta(1,2,3)$ 的三个顶点及三边中点为插值节点 (参看图 4.3(b)). 仍用待定系数法, 得

$$P_2(x,y)=\sum_{i=1}^{3}[L_i(2L_i-1)u_i+4L_jL_ku_{3+i}], \tag{6.4.9}$$

其中 u 的下标 4, 5, 6 依次是边 23, 31, 12 的中点, $L_j=L_{i+1}, L_k=L_{i+2}, L_4=L_1, L_5=L_2, L_6=L_3$.

还可以构造三次及高次的 Lagrange 型公式, 但常用的是一次及二次插值公式.

容易证明, 由 m 次 Lagrange 型插值公式生成的试探函数属于 C, 但不属于 C^1, 因此只能有 $U_h\subset H^1$.

从插值公式 (6.4.8), (6.4.9) 可直接得出基函数. 例如从公式 (6.4.8) 知道, L_1 是相应于节点 1 的基函数在单元 $\Delta(1,2,3)$ 上的限制.

4.3　Hermite 型公式

考虑三次 Hermite 插值公式. 三次多项式总可写成形式:

$$P_3(x,y) = c_1L_1^3 + c_2L_2^3 + c_3L_3^3 + c_4L_1^2L_2 + c_5L_1^2L_3 + c_6L_2^2L_1 + c_7L_2^2L_3 + c_8L_3^2L_1 + c_9L_3^2L_2 + c_{10}L_1L_2L_3,$$

其中 10 个系数按下述 10 个插值条件确定: 在三个顶点上, $P_3(x,y)$ 及其一阶导数分别取 u 的相应值 $u_i, (u_x)_i, (u_y)_i (i=1,2,3)$. 此外, 在 $\Delta(1,2,3)$ 的重心处取 u 的函数值 u_0 (0 表示重心, 参看图 4.4).

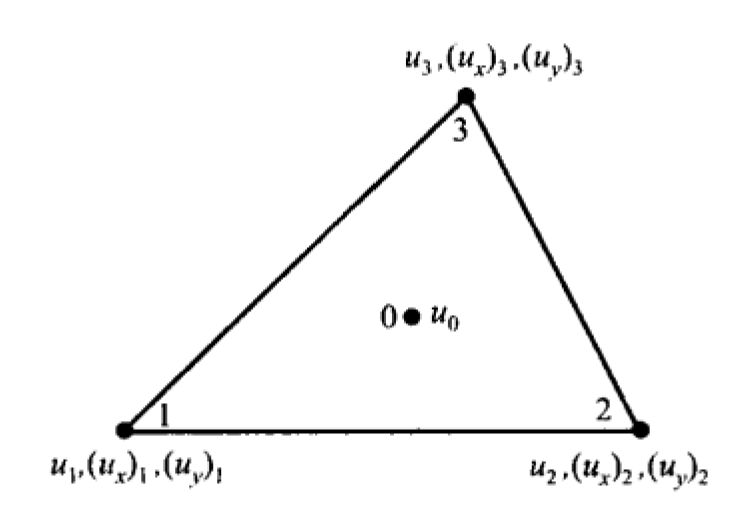


图 4.4

用待定系数法, 可求出此多项式为

(6.4.10)

$$\begin{cases} P_3(x,y) = \alpha_0^3(x,y)u_0 + \sum_{i=1}^{3}[\alpha_i^{(3)}(x,y)u_i + \beta_i^{(3)}(x,y)(u_x)_i + \gamma_i^{(3)}(x,y)(u_y)_i], \\ \alpha_0^{(3)}(x,y) = 27L_1L_2L_3, \\ \alpha_i^{(3)}(x,y) = L_i^3 + 3L_i^2(L_j+L_k) - 7L_iL_jL_k, \\ \beta_i^{(3)}(x,y) = (x_j-x_i)(L_i^2L_j - L_iL_jL_k) + (x_k-x_i)(L_i^2L_k - L_iL_jL_k), \\ \gamma_i^{(3)}(x,y) = (y_j-y_i)(L_i^2L_j - L_iL_jL_k) + (y_k-y_i)(L_i^2L_k - L_iL_jL_k), \end{cases}$$

这里 j,k 应如此选取, 使得 i,j,k 成逆时针方向.

一旦有了插值公式 (6.4.10), 基函数就自然造出来了. 实际上, 和节点 1 相应的基函数在 $\Delta(1,2,3)$ 上的限制就是 $\alpha_1^{(3)}(x,y), \beta_1^{(3)}(x,y), \gamma_1^{(3)}(x,y)$.$\Delta(1,2,3)$ 是任一以 1 为顶点的三角单元, 所以实际上得到了和节点 1 相应的基函数. 同样, 和重心 0 相应的基函数在单元内部等于 $\alpha_0^{(3)}(x,y)$, 在外部等于 0.

给定了一组广义坐标后 (包括三角单元顶点的函数值, 两个一阶偏导数及单元重心的函数值), 由公式 (6.4.10) (让 $\Delta(1,2,3)$ 取遍一切单元) 确定出整个剖分上的试探函数 $u_h(x,y)$ (下标 h 表示一切单元的最大直径). 一切可能的试探函

数构成试探函数空间 U_h. 设 K_1, K_2 是两相邻单元, l 为其公共边, u_h^1 和 u_h^2 分别表示 u_h 从 K_1 和 K_2 延拓到 l 上的一元三次多项式. 因为 u_h^1 和 u_h^2 在 l 的两端取相同函数值及沿 l 方向的导数值 (它们由偏导数 u_x, u_y 唯一确定), 故 u_h^1 和 u_h^2 在 l 上恒等, 这说明 $u_h \in C$, 因而 $U_h \subset H^1$. 其次, 考察它们的一阶导数. 显然 u_h^1 和 u_h^2 沿 l 方向的导数恒等, 但沿 l 的法向导数不一定相等. 因为法向导数是 l 上的一元二次多项式, 二者仅在端点相等, 不能完全确定该二次多项式. 因此 u_h 一般不属于 C^1, 从而也不能要求 U_h 属于 H^2.

应当指出, 重心点的方程只包含重心值及它所属单元三个顶点的广义坐标, 所以求解有限元方程时, 可先消去重心处的广义坐标, 降低方程组的阶, 然后再解这个阶数较低的方程组.

习 题

1. 证明积分公式 (6.4.6).

2. 设 l 是 xy 平面上的直线, 方程为 $ax+by+c=0$ $(a^2+b^2=1)$, $\boldsymbol{n}$ 是 l 的单位法向量. 若多项式 p 具性质

$$\left.\frac{\mathrm{d}^i p}{\mathrm{d}\boldsymbol{n}^i}\right|_l = 0, \quad 0 \leqslant i \leqslant k,$$

则 p 可用 $(ax+by+c)^{k+1}$ 整除, 即有多项式 $q(x,y)$ 使 $p(x,y)=(ax+by+c)^{k+1}q(x,y)$.

*§5 曲边元和等参变换

如果区域 G 的边界 $\varGamma$ 过于复杂, 以致用折线逼近的几何误差太大, 就需采取分段高次曲线逼近, 这时将出现曲边单元. §3 曾讨论过平行坐标轴的矩形剖分及矩形单元的形状函数, 这类剖分对规则区域才是方便的; 否则可采取任意四边形剖分, 它和三角剖分一样有很大的灵活性. 对于上述曲边单元、四边形单元, 如何构造单元形状函数? 这是本节要讨论的问题.

前两节构造单元形状函数时, 总是用一个变换 —— 仿射变换, 把任意单元 K 变到 $\xi\eta$ 平面上的 "参考元" E, E 是一个单位正方形或单位直角三角形 (参看图 3.1 和图 4.2), 其几何形状非常简单, 易于构造单元形状函数. 其实 K 也是参考元 E 在仿射变换下的映像. 不同的仿射变换, 便得到不同的单元及其形状函数. 所以, G 的任一网格剖分及试探函数, 可以看作是参考元 E 及其形状函数在一族仿射变换下得到的.

例 5.1 设参考单元 E 为单位正方形: $0 \leqslant \xi \leqslant 1, 0 \leqslant \eta \leqslant 1$ (参看图 5.1(a)). 由 (6.3.4), E 上的双线性函数为

$$p_E(\xi,\eta) = (1-\xi)(1-\eta)u_1 + \xi(1-\eta)u_2 + (1-\xi)\eta u_3 + \xi\eta u_4. \tag{6.5.1}$$

在仿射变换 (参看 (6.3.1))

$$
(6.5.2)\qquad \begin{cases} x=(1-\xi)(1-\eta)x_1+\xi(1-\eta)x_2+(1-\xi)\eta x_3+\xi\eta x_4, \\ y=(1-\xi)(1-\eta)y_1+\xi(1-\eta)y_2+(1-\xi)\eta y_3+\xi\eta y_4 \end{cases}
$$

$(x_3=x_1, x_4=x_2, y_2=y_1, y_4=y_3)$ 下, E 变到以 $(x_1,y_1),(x_2,y_2),(x_3,y_3),(x_4,y_4)$ 为顶点的矩形 R (参看图 5.1(b)), 而 $p_E(\xi,\eta)$ 变成 R 上的双线性函数 $p_R(x,y)$.

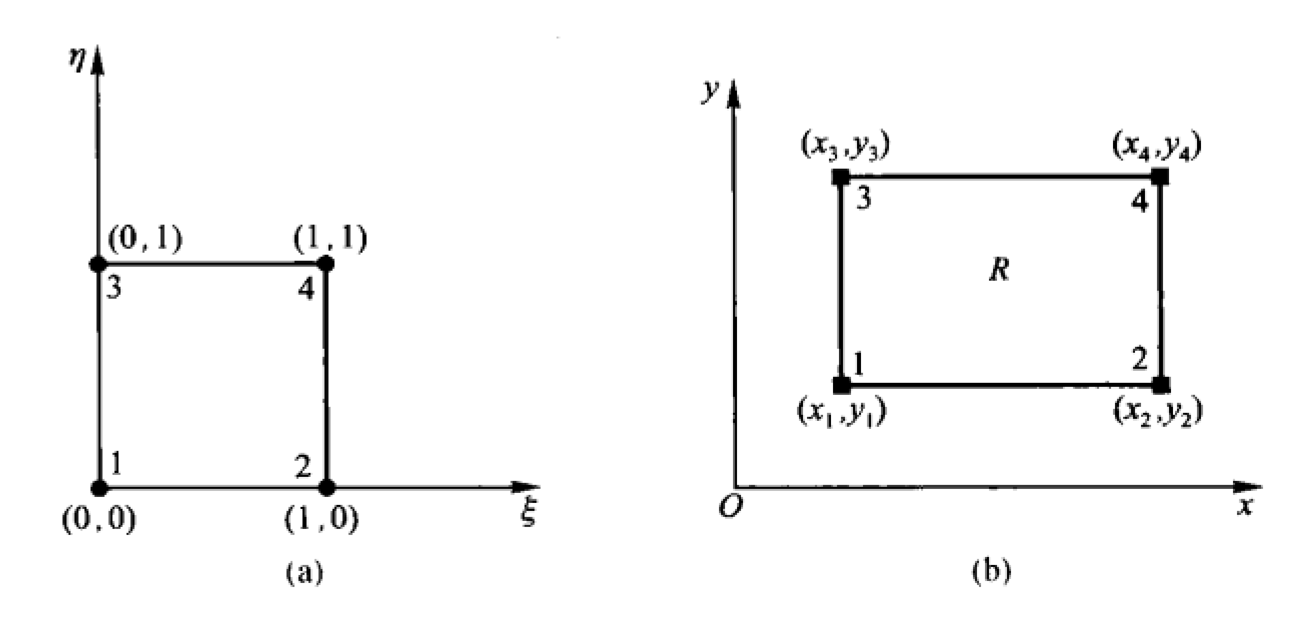


图 5.1

现在我们面临的是一般单元 K (四边形单元、曲边单元), 只限于仿射变换是不够的. 因此需要考虑更一般的可逆连续变换:

$$
(6.5.3)\qquad F=F_K(\xi,\eta)=(x(\xi,\eta),\ y(\xi,\eta)),
$$

其中 $(\xi,\eta)\in E,(x,y)\in K$. 变换 (6.5.3) 和单元 K 有关, 不同的 K 对应不同的变换. 也可反过来说, 不同的变换对应不同的单元 K. 变换 (6.5.3) 应满足下列要求: 第一, 具有必要的光滑性. 通常取它为 ξ,η 的多项式, 所以光滑性条件恒满足. 第二, F_K 应是 E 到 K 的一对一的变换, 就是说, F_K 的 Jacobi 行列式

$$
(6.5.4)\qquad J(\xi,\eta)=\frac{\partial(x,y)}{\partial(\xi,\eta)}\neq 0,\quad (\xi,\eta)\in E.
$$

应指出的是, E 上的形状函数 $p_E(\xi,\eta)$ 虽然是多项式, 但是通过变换 F_K 的逆变换消去 ξ,η 后就不一定是多项式了, 它可能是有理函数, 也可能是无理函数. 好在我们并不需要消去 ξ,η 得到以 x,y 表示的形状函数, 因为形成有限元方程时遇到的积分可通过变换 (6.5.3) 化为 E 上对 ξ,η 的积分 (见本章 §6).

构造变换 (6.5.3) 的方法很多, 应用中最重要的一种是取它和形状函数具同样形式, 这就是所谓 "**等参变换**". 例如变换 (6.5.2) 和形状函数 (6.5.1) 都是双线性函数. 使用等参变换也是有限元法的一个技巧.

例 5.2 (任意四边形单元) 如图 5.2, E 是 $\xi\eta$ 平面上的单位正方形; K 是 xy 平面上任一四边形, 顶点是 $(x_1,y_1),(x_2,y_2),(x_3,y_3),(x_4,y_4)$. 我们知道, E 上最简单的形状函数是双线性函数 (6.5.1). 显然 (6.5.2) 就是把 E 变到 K 的双线性等参变换. 此时变换 (6.5.3) 形如 (6.5.2), 或者等价地写成

$$(6.5.5)\qquad \begin{cases} x(\xi,\eta)=x_1+(x_2-x_1)\xi+(x_3-x_1)\eta+(x_4-x_3-x_2+x_1)\xi\eta, \\ y(\xi,\eta)=y_1+(y_2-y_1)\xi+(y_3-y_1)\eta+(y_4-y_3-y_2+y_1)\xi\eta. \end{cases}$$

显然, 变换 (6.5.5) 把 E 的每条边仿射变换到 K 的对应边, 把 E 的内部变到 K 的内部.

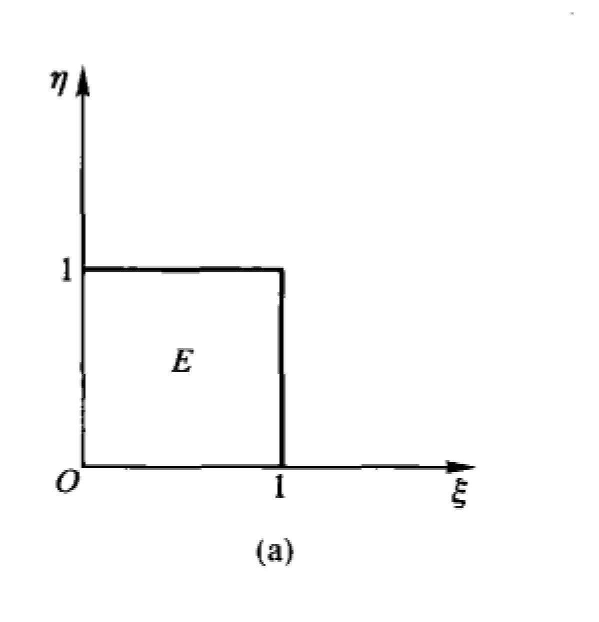
为了检验变换 (6.5.5) 是一对一的, 计算 Jacobi 行列式

$$J=\begin{vmatrix} x_2-x_1+A\eta & x_3-x_1+A\xi \\ y_2-y_1+B\eta & y_3-y_1+B\xi \end{vmatrix},$$

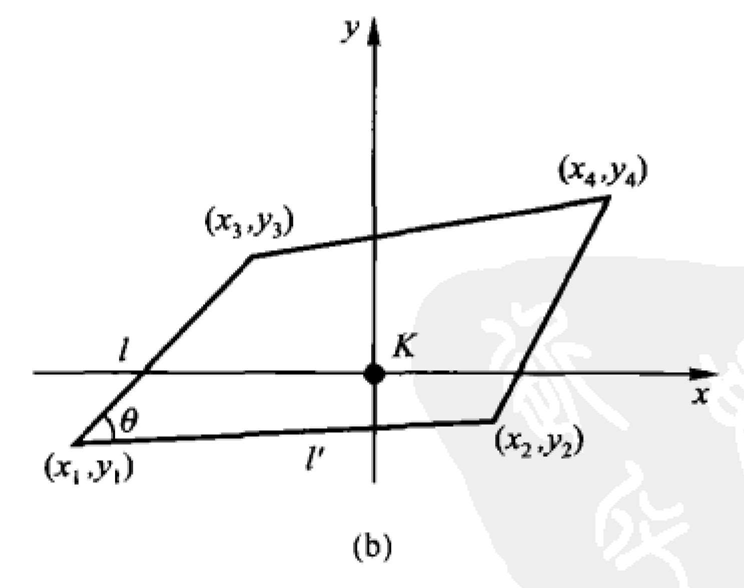
其中 $A=x_4-x_3-x_2+x_1, B=y_4-y_3-y_2+y_1$. 展开后, J 是 ξ,η 的双线性函数. 故只需检验 J 在四个顶点的值有相同符号即可. 于 $\xi=0,\eta=0$,

$$J(0,0)=(x_2-x_1)\ (y_3-y_1)-(y_2-y_1)\ (x_3-x_1)=ll'\sin\theta,$$

长度 l,l' 和角 θ 如图 5.2(b) 所示. 当 $0<\theta<\pi$ 时, $J(0,0)>0$. 同样, 其他三个内角小于 π 时, J 在 $(1,0),(0,1),(1,1)$ 也大于 0. 总之, J 于 E 恒大于 0 的充分必要条件是 K 为凸四边形.



(a)



(b)

图 5.2

注意 (6.5.5) 的逆变换是无理函数, 所以 E 上的形状函数尽管是简单的 (双线性函数), 但对应到 K 上则是 x,y 的无理函数. 好在我们不必通过 x,y 把 E 上的形状函数以显式表示出来 (参看 §6).

变换 (6.5.5) 很重要, 通过它还可在 K 上构造其他形状函数. 例如, 对 E 上的双二次函数 (6.3.8), 通过 (6.5.5) 相应地可确定出 K 上一精度更高的形状函数. 此时变换 (6.5.5) 的阶数比形状函数 (6.3.8) 的阶数低, 所以相对双二次函数来说, (6.5.5) 是**次参变换**.

例 5.3 (曲边三角形单元) 如图 5.3(a), $K=\widetilde{\Delta}(1,2,3)$ 是 xy 平面上一个曲边三角形单元, Γ_{23} 是外边界 Γ 的一段曲线弧, 其余两边 $\overline{12},\overline{13}$ 是直线段. 设顶点 $1,2,3$ 的坐标为 $(x_1,y_1),(x_2,y_2),(x_3,y_3)$. 在构造 $\widetilde{\Delta}(1,2,3)$ 上的形状函数时, 先通过一个变换 (6.5.3) 把它变到 $\xi\eta$ 平面上的参考单元 E (参看图 5.3(c)).

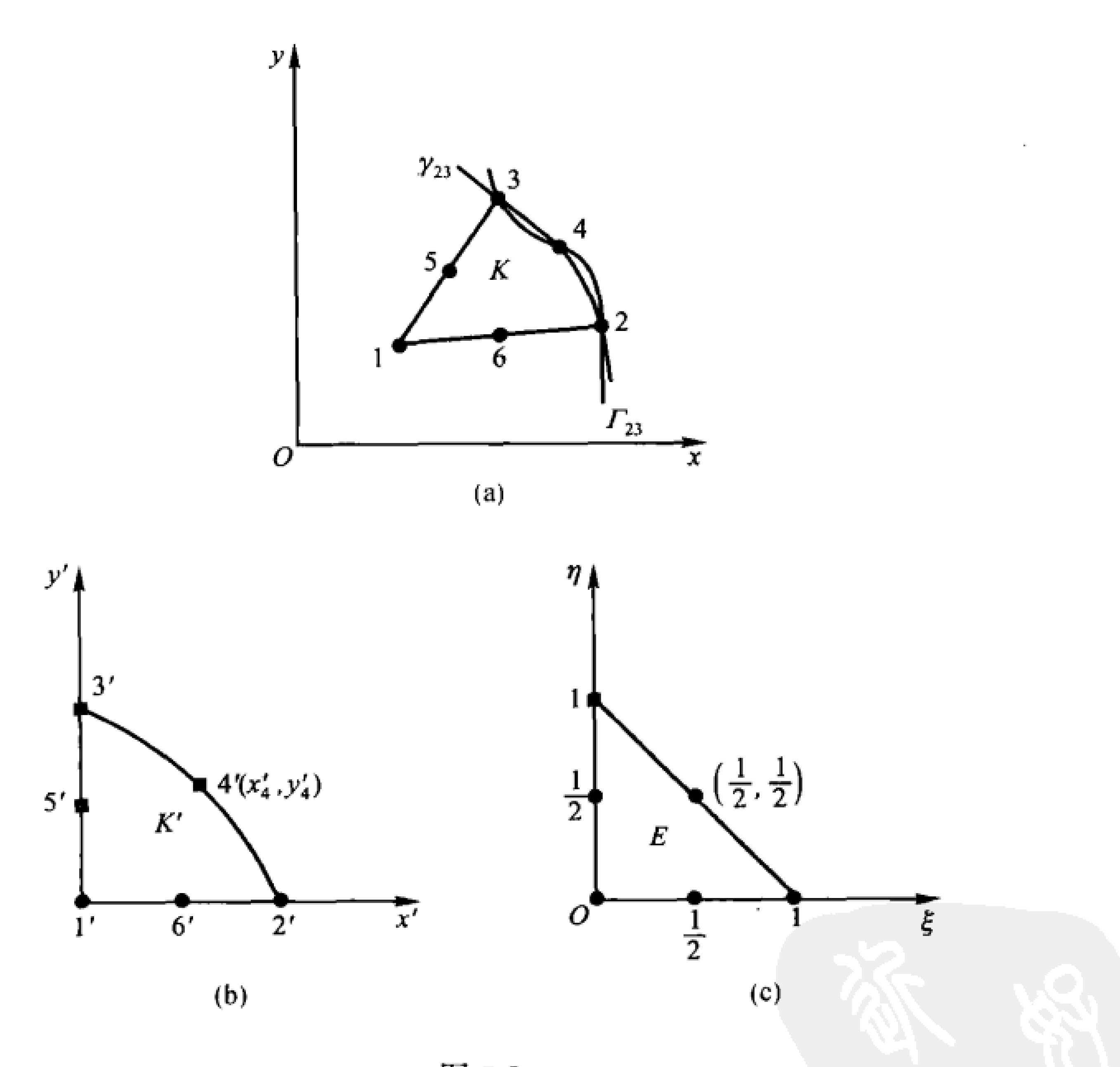


图 5.3

为此, 我们用一代数曲线 γ_{23} 插值逼近 Γ_{23}. 例如, 取 γ_{23} 为二次曲线. 为此, 于 Γ_{23} 内部取一点 4, 其坐标为 (x_4,y_4). 作仿射变换 G_K:

$$(x,y)=G_K(x',y')=(g_1(x',y'),g_2(x',y')), \tag{6.5.6}$$

把曲边三角形 $K=\widetilde{\Delta}(1,2,3)$ 变到 $x'y'$ 平面上的一曲边三角形 $K'=\widetilde{\Delta}(1',2',3')$, 使顶点 $(x_1,y_1),(x_2,y_2),(x_3,y_3)$ 依次对应 K' 的顶点 $1',2',3'$, 其坐标分别为 $(0,0)$,

$(1,0),(0,1)$ (参看图 5.3(b)). 显然

$$
(6.5.7)\qquad \begin{cases} x = g_1(x',y') = x_1 + (x_2 - x_1)x' + (x_3 - x_1)y', \\ y = g_2(x',y') = y_1 + (y_2 - y_1)x' + (y_3 - y_1)y'. \end{cases}
$$

设 5,6 分别为 $\overline{13},\overline{12}$ 的中点. 经过仿射变换 (6.5.7), 点 5,6 变到 $x'y'$ 平面上是线段 $\overline{1'3'},\overline{1'2'}$ 的中点 $5',6'$; 点 4 变成曲线 $\widetilde{2'3'}$ 上一点 $4'$, 其坐标 (x_4',y_4') 由 (6.5.7) 决定. 现在构造一个从 K' 到 E 的二次变换 $H_{K'}$

$$
(6.5.8)\qquad (x',y') = H_{K'}(\xi,\eta) = (h_1(\xi,\eta),h_2(\xi,\eta)),
$$

使

$$
\begin{gathered} 1' \to (0,0), \quad 2' \to (1,0), \quad 3' \to (0,1), \\ 5' \to \left(0,\frac{1}{2}\right), \quad 6' \to \left(\frac{1}{2},0\right), \quad 4' \to \left(\frac{1}{2},\frac{1}{2}\right). \end{gathered}
$$

用待定系数法, 得

$$
(6.5.9)\qquad \begin{cases} x' = h_1(\xi,\eta) = \xi + (4x_4' - 2)\xi\eta, \\ y' = h_2(\xi,\eta) = \eta + (4y_4' - 2)\xi\eta. \end{cases}
$$

变换 (6.5.9) 的 Jacobi 行列式

$$
J = \begin{vmatrix} 1 + (4x_4' - 2)\eta & (4x_4' - 2)\xi \\ (4y_4' - 2)\eta & 1 + (4y_4' - 2)\xi \end{vmatrix} = 1 + (4y_4' - 2)\xi + (4x_4' - 2)\eta
$$

是线性函数, 在 $(0,0)$ 等于 1. 为使其在 E 上不等于 0, 只要它在另二顶点 $(1,0),(0,1)$ 大于 0, 即

$$
(6.5.10)\qquad \begin{cases} (4x_4' - 2) > -1 \ \text{或}\ x_4' > \dfrac{1}{4}, \\ (4y_4' - 2) > -1 \ \text{或}\ y_4' > \dfrac{1}{4}. \end{cases}
$$

因此, 在曲边 Γ_{23} 上的点 4 应如此选择, 使它在仿射变换 G_K 下的映像 (x_4',y_4') 满足 (6.5.10).

将 $G_K,H_{K'}$ 结合, 就得到所需的变换 (6.5.3), 其中

$$
(6.5.11)\qquad \begin{cases} x = x(\xi,\eta) = x_1 + (x_2 - x_1)[\xi + 2(2x_4' - 1)\xi\eta] + \\ \qquad\qquad (x_3 - x_1)[\eta + 2(2y_4' - 1)\xi\eta], \\ y = y(\xi,\eta) = y_1 + (y_2 - y_1)[\xi + 2(2x_4' - 1)\xi\eta] + \\ \qquad\qquad (y_3 - y_1)[\eta + 2(2y_4' - 1)\xi\eta]. \end{cases}
$$

它把 E 的两个直角边变到 K 的两个直边, E 的斜边 $\xi+\eta=1$ 变成通过 2,3,4 的二次曲线 γ_{23}. 若用 $\widetilde{K}$ 表示 xy 平面上由直线段 $\overline{12},\overline{13}$ 和曲线 γ_{23} 围成的曲边三角形, 则 (6.5.11) 是将 E 变到 $\widetilde{K}$ 的一对一变换. $\widetilde{K}$ 和 K 的唯一差别是: 用二次曲线 γ_{23} 代替曲边 Γ_{23}, 二者于点 2,4,3 重合 (参看图 5.3(a)). 特别当 $\widetilde{\Delta}(1,2,3)$ 是直边三角形时, (6.5.11) 退化为仿射变换.

通过变换 (6.5.11), 就不难构造 K 上的形状函数. 实际上, 参考单元 E 上任一形状函数 $p_E(\xi,\eta)$, 通过变换 (6.5.11) 便得出 K 上的形状函数 $p_K(x,y)$. 如前指出, 我们无需 $p_K(x,y)$ 的显式表达式, 只知道确定它的 $p_E(\xi,\eta)$ 和关系式 (6.5.11) 就够了.

§6 二阶椭圆方程的有限元法

确定了网络剖分和单元形状函数后, 试探函数空间 U_h 也就定了. 本节讨论二阶椭圆方程的有限元法.

6.1 有限元方程的形成

如本章 §1 所述, 有两种途径形成有限元方程. 假定已按一定次序分别把节点和单元编号. 第一种是工程界流行的方法, 它从 Ritz 法及单元形状函数出发, 先形成单元刚度矩阵 (其阶数等于单元自由度), 然后由单元刚度矩阵叠加成总刚度矩阵. 叠加的规则与一维情形完全相同 (参看 §1).

第二种方法是从 Galerkin 法及节点基函数出发 (一个节点有几个广义坐标, 就有几个基函数), 直接形成有限元方程. 例如, 设全部基函数为

$$\varphi_1,\cdots,\varphi_n,$$

边值问题的变分形式为: 求 $u\in H_E^1$ 使

$$a(u,v)=(f,v),\qquad 对任意\ v\in H_E^1, \tag{6.6.1}$$

则有限元方程为

$$\sum_{i=1}^{n} a(\varphi_i,\varphi_j)u_i=(f,\varphi_j),\quad j=1,\cdots,n \tag{6.6.2}$$

(参看第五章 §5). 在方程左端系数中, 只有当 φ_i,φ_j 对应的节点相邻时才不为 0. 又因为 φ_i,φ_j 都是由它在单元上的限制 “拼凑” 而成的, 所以实际计算时归结到单元上的积分, 这和第一种方法类似. 不同的是, 前者从单元出发 (形成单元刚度矩阵), 然后形成有限元方程 (形成总刚度矩阵); 后者则直接按节点形成方程.

显然, 方程组 (6.6.2) 的系数矩阵稀疏. 若注意到双线性形式 $a(u,v)$ 对称正定, 则系数矩阵也对称正定 (参看第五章 §5). 由此即可结论: 有限元方程组的解存在唯一. 方程组的这些特性, 为我们选择解法提供了依据.

6.2 矩阵元素的计算

形成有限元方程的主要工作量是计算大量的积分, 通过变量代换 —— 仿射变换或等参变换, 可把它们化成 $\xi\eta$ 平面上参考单元 E(单位矩形, 直角三角形) 上的积分. 例如边值问题

$$
(6.6.3)\qquad \begin{cases} -\nabla(a(x,y)\nabla u(x,y)) = f(x,y), & (x,y)\in G, \\ u|_\Gamma = 0, \end{cases}
$$

与它相应的双线性形式为

$$
\begin{aligned}
a(u,v) &= \iint_G a(x,y)[u_x v_x + u_y v_y]\mathrm{d}x\mathrm{d}y \\
&= \sum_K \iint_K a(x,y)[u_x v_x + u_y v_y]\mathrm{d}x\mathrm{d}y,
\end{aligned}
$$

其中 K 是网格剖分单元. 利用变换 (6.5.3) 有

$$
\begin{aligned}
&\iint_K a(x,y)[u_x v_x + u_y v_y]\mathrm{d}x\mathrm{d}y \\
=&\iint_E a\circ F_K \left[\left((u\circ F_K)'_\xi\cdot\frac{\partial\xi}{\partial x} + (u\circ F_K)'_\eta\cdot\frac{\partial\eta}{\partial x}\right)\left((v\circ F_K)'_\xi\cdot\frac{\partial\xi}{\partial x}+\right.\right. \\
&\left.(v\circ F_K)'_\eta\cdot\frac{\partial\eta}{\partial x}\right) + \left((u\circ F_K)'_\xi\cdot\frac{\partial\xi}{\partial y}+(u\circ F_K)'_\eta\cdot\frac{\partial\eta}{\partial y}\right)\cdot \\
&\left.\left((v\circ F_K)'_\xi\cdot\frac{\partial\xi}{\partial y} + (v\circ F_K)'_\eta\cdot\frac{\partial\eta}{\partial y}\right)\right] J(\xi,\eta)\mathrm{d}\xi\mathrm{d}\eta,
\end{aligned}
$$

其中 $a\circ F_K$ 表示 $a(x(\xi,\eta),y(\xi,\eta))$ ($u\circ F_K, v\circ F_K$ 的意思类似). 而

$$
\frac{\partial\xi}{\partial x} = \frac{\partial y}{\partial\eta}\Big/J,\quad \frac{\partial\eta}{\partial x} = -\frac{\partial y}{\partial\xi}\Big/J,
$$

$$
\frac{\partial\eta}{\partial y} = \frac{\partial x}{\partial\xi}\Big/J,\quad \frac{\partial\xi}{\partial y} = -\frac{\partial x}{\partial\eta}\Big/J,
$$

J 是变换 $x=x(\xi,\eta), y=y(\xi,\eta)$ 的 Jacobi 行列式. 于是

$$
\begin{aligned}
(6.6.4)\qquad & \iint_K \alpha(u_x v_x + u_y v_y)\mathrm{d}x\mathrm{d}y \\
= & \iint_E \alpha\circ F_K \left[\left((u\circ F_K)'_\xi\cdot\frac{\partial y}{\partial\eta} - (u\circ F_K)'_\eta\cdot\frac{\partial y}{\partial\xi}\right)\cdot\right. \\
& \left((v\circ F_K)'_\xi\cdot\frac{\partial y}{\partial\eta} - (v\circ F_K)'_\eta\cdot\frac{\partial y}{\partial\xi}\right) + \\
& \left((u\circ F_K)'_\eta\cdot\frac{\partial x}{\partial\xi} - (u\circ F_K)'_\xi\cdot\frac{\partial x}{\partial\eta}\right)\cdot \\
& \left.\left((v\circ F_K)'_\eta\cdot\frac{\partial x}{\partial\xi} - (v\circ F_K)'_\xi\cdot\frac{\partial x}{\partial\eta}\right)\right] J^{-1}(\xi,\eta)\mathrm{d}\xi\mathrm{d}\eta.
\end{aligned}
$$

特别取 u, v 为基函数 φ_i, φ_j, 则

$$
\begin{aligned}
(6.6.5)\qquad a(\varphi_i,\varphi_j) = & \sum_K \iint_E \alpha\circ F_K \left[\left((\varphi_i\circ F_K)'_\xi\cdot\frac{\partial y}{\partial\eta} - (\varphi_i\circ F_K)'_\eta\cdot\frac{\partial y}{\partial\xi}\right)\cdot\right. \\
& \left((\varphi_j\circ F_K)'_\xi\cdot\frac{\partial y}{\partial\eta} - (\varphi_j\circ F_K)'_\eta\cdot\frac{\partial y}{\partial\xi}\right) + \\
& \left((\varphi_i\circ F_K)'_\eta\cdot\frac{\partial x}{\partial\xi} - (\varphi_i\circ F_K)'_\xi\cdot\frac{\partial x}{\partial\eta}\right)\cdot \\
& \left.\left((\varphi_j\circ F_K)'_\eta\cdot\frac{\partial x}{\partial\xi} - (\varphi_j\circ F_K)'_\xi\cdot\frac{\partial x}{\partial\eta}\right)\right] J^{-1}(\xi,\eta)\mathrm{d}\xi\mathrm{d}\eta.
\end{aligned}
$$

其中 $\varphi_i\circ F_K, \varphi_j\circ F_K$ 是参考单元 E 上的基函数, 都是 ξ,η 的多项式, 而 $J^{-1}(\xi,\eta)$ 是 ξ,η 的有理函数.

如果 $\alpha\circ F_K, \varphi_i\circ F_K$ 和 $\varphi_j\circ F_K$ 是低次多项式,F_K 是仿射变换, 则积分 (6.6.5) 容易计算 (公式 (6.4.6) 对计算这类积分很有用). 在一般情形, 需要运用数值积分公式, 通常用 Gauss 求积公式.

为了便于读者查阅, 列出三角形 $\Delta(1,2,3)$ 上的 Gauss 求积公式

$$
\iint_\Delta f\mathrm{d}x\mathrm{d}y = S\sum_i W_i f(L_1^{(i)}, L_2^{(i)}, L_3^{(i)}),
$$

其中 S 是 Δ 的面积,$(L_1^{(i)}, L_2^{(i)}, L_3^{(i)})$ 是求积节点的面积坐标,W_i 是求积系数. 在本节的表 6.1 中, 列出了求积公式系数和节点坐标. 公式关于面积坐标 (L_1, L_2, L_3) 是对称分布的. 若表中出现的 L_1, L_2, L_3 互异, 则经过置换应有六个求积节点. 若表中出现的 L_1, L_2, L_3 有两个互异, 则经置换后应有三个求积节点. 点 $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$ 是重心, 若取作求积节点, 则只出现一次 (见 [29]).

6.3 边值条件的处理

对于剖分的边界节点, 要根据边值条件的类型作适当处理, 处理方法和一维情形类似 (见 §1).

考虑 Poisson 方程

$$-\Delta u = f(x,y), \quad (x,y) \in G. \tag{6.6.6}$$

G 是 xy 平面的一个有界域, 其边界 Γ 是分段光滑的简单闭曲线. 在 Γ 上给出下列边值条件之一:

$$(6.6.6)_1 \qquad u|_\Gamma = \alpha(x,y) \qquad (\text{第一边值条件}),$$

$$(6.6.6)_2 \qquad \left.\frac{\partial u}{\partial n}\right|_\Gamma = \beta(x,y) \qquad (\text{第二边值条件}),$$

$$(6.6.6)_3 \qquad \left.\frac{\partial u}{\partial n} + ku\right|_\Gamma = \gamma(x,y) \qquad (\text{第三边值条件}),$$

其中 $f(x,y), \alpha(x,y), \beta(x,y), \gamma(x,y)$ 及 $k(x,y)$ 都是给定的连续函数. 边值条件也可以这样给: 在 Γ 的一部分满足一种边值条件, 其余部分满足另一种边值条件.

不妨假定 G 是多边形域, Γ 是 G 的外边界. 对 G 作剖分, 比如是三角剖分, 如图 6.1 所示. 图中 1,2,4,7,8,9,10 是边界点, 3,5,6 是内点. 同内点一样, 在界点也引进基函数, 引进的方法及计算公式和内点完全一样. 与任一节点 i 相应的基函数之支集是一切以 i 为顶点的三角单元. 比如与内点 3 相应的基函数之支集是单元 $\Delta(3,1,2), \Delta(3,2,6), \Delta(3,6,5), \Delta(3,5,4)$ 及 $\Delta(3,4,1)$; 与界点 1 相应的基函数之支集是单元 $\Delta(1,2,3)$ 及 $\Delta(1,3,4)$. 至于与节点 i 相应的基函数的个数, 要根据节点 i 的广义坐标的数目而定. 比如线性元, 每一点 i 的广义坐标数是 1, 即函数值 u_i, 此时与 i 相应的基函数只有一个, 记为 φ_i. 若取 Hermite 三次元为形状函数, 则节点包含 10 个顶点 $1,2,\cdots,10$ 和 11 个单元重心, 编号为 $11,12,\cdots,21$. 对每一顶点 $i: 1 \leqslant i \leqslant 10$, 有三个广义坐标, 即函数值 u_i 及其偏导数 $(u_x)_i, (u_y)_i$, 相应的基函数有三个, 设为 $\varphi_i, (\varphi_x)_i, (\varphi_y)_i$. 对每一单元重心 $i: 11 \leqslant i \leqslant 21$, 有一个广义坐标 u_i 和一个基函数 φ_i.

先讨论第一边值条件. 若形状函数是线性函数, 则就内点写出的方程 (6.6.2) 形如

$$\sum_{i=1}^{10} a(\varphi_i, \varphi_j) u_i = (f, \varphi_j), \quad j = 3,5,6. \tag{6.6.7}$$

然后用边值 $u_i = \alpha_i = \alpha(x_i, y_i)$ $(i = 1,2,4,7,8,9,10)$ 代入左端相应项, 并移至右端, 就得出有限元方程 (三元一次方程). 若形状函数是 Hermite 三次函数, 则方

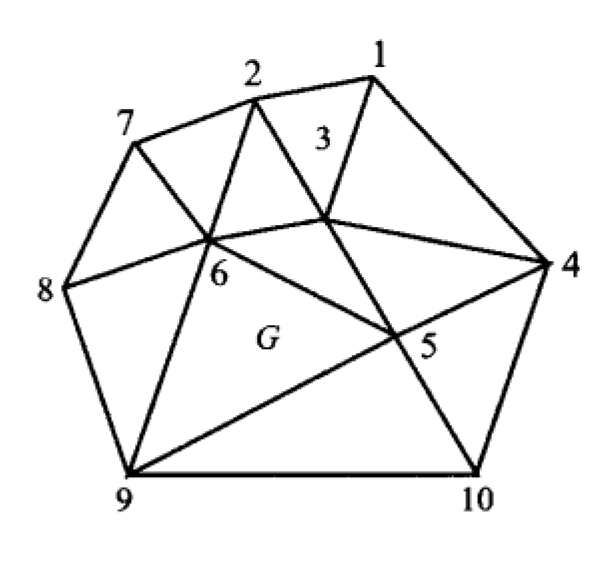


图 6.1

程 (6.6.2) 的具体形式是

$$(6.6.8)_1 \quad \sum_{i=1}^{21} a(\varphi_i,\varphi_j)u_i + \sum_{i=1}^{10}[a((\varphi_x)_i,\varphi_j)\ (u_x)_i + a((\varphi_y)_i,\varphi_j)(u_y)_i] = (f,\varphi_j), \quad j=3,5,6,11,12,\cdots,21,$$

$$(6.6.8)_2 \quad \sum_{i=1}^{21} a(\varphi_i,(\varphi_x)_j)u_i + \sum_{i=1}^{10}[a((\varphi_x)_i,(\varphi_x)_j)\ (u_x)_i + a((\varphi_y)_i,(\varphi_x)_j)(u_y)_i] = (f,(\varphi_x)_j), \quad j=1,2,\cdots,10,$$

$$(6.6.8)_3 \quad \sum_{i=1}^{21} a(\varphi_i,(\varphi_y)_j)u_i + \sum_{i=1}^{10}[a((\varphi_x)_i,(\varphi_y)_j)\ (u_x)_i + a((\varphi_y)_i,(\varphi_y)_j)(u_y)_i] = (f,(\varphi_y)_j), \quad j=1,2,\cdots,10.$$

对于内点 $j=3,5,6$, 上述三个方程都保留. 对于重心点 $j=11,12,\cdots,21$, 方程 $(6.6.8)_1$ 保留. 对于界点 $j=1,2,4,7,8,9,10$, 方程 $(6.6.8)_2$ 和 $(6.6.8)_3$ 保留. 然后就一切界点 i 用边值 $u_i=\alpha_i=\alpha(x_i,y_i)$ 代入方程左端相应项, 并移到右端, 为右端提供贡献. 这就得到我们所要的方程. 每个内点有三个方程, 边界点有两个方程, 重心点有一个方程, 共计 $3\times3+7\times2+11\times1=34$ 个方程. 对其他类型的形状函数, 处理方法类似.

现在讨论第二、第三边值条件. 因为它们是自然边值条件, 可像内点一样在边界点形成有限元方程, 只不过双线性形式 $a(u,v)$ 因边值条件不同需要修改. 例如对 Poisson 方程 (6.6.6) 和边值条件 $(6.6.6)_2$, $(6.6.8)_3$, 相应的双线性形式分别为 (参看第五章 §4)

$$(6.6.9)_1 \quad a(u,v)=\iint\limits_G \left[\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right]\mathrm{d}x\mathrm{d}y,$$

$$(6.6.9)_2 \quad a(u,v)=\iint\limits_G \left[\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right]\mathrm{d}x\mathrm{d}y+\int_\Gamma kuv\mathrm{d}s.$$

有限元方程右端分别修改为

$$\iint\limits_G fv\mathrm{d}x\mathrm{d}y + \int_\Gamma \beta v\mathrm{d}s$$

和

$$\iint\limits_G fv\mathrm{d}x\mathrm{d}y + \int_\Gamma \gamma v\mathrm{d}s.$$

若注意到基函数支集的局限性, 则知曲线积分只对界点及与界点相邻节点上的有限元方程有影响.

最后讨论混合边值条件: 在 Γ 的一部分 Γ_1 给第一边值, 在其余部分 Γ_2 给第二或第三边值. 这时可用前述方法分别处理 Γ_1 和 Γ_2. Γ_1, Γ_2 的交界点应取作剖分节点, 把它当作第一边值点处理.

6.4 举例: Poisson 方程的有限元法

至今, 我们已讨论了有限元计算的有关问题. 这些问题是: 把边值问题化成变分形式, 对求解域作网格剖分, 构造基函数 (或单元形状函数), 形成有限元方程.

现在, 我们以 Poisson 方程第一边值问题 (6.6.6), $(6.6.6)_1$ 为例说明用有限元法解题的主要过程.

第一, 把边值问题 (6.6.6), $(6.6.6)_1$ 化成等价的变分形式: 求 $u \in H^1(G)$, $u|_\Gamma = \alpha$, 使

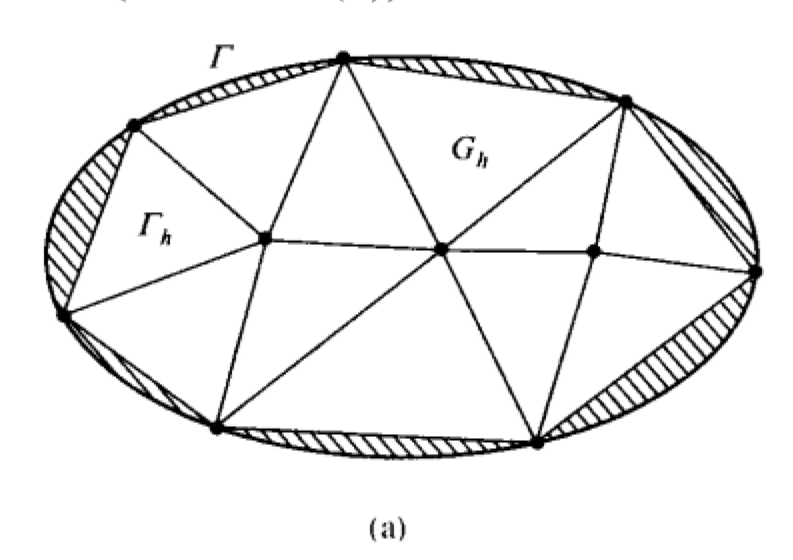
$$a(u,v) = f(u,v), \quad \text{对任意 } v \in H_0^1(G), \tag{6.6.10}$$

其中

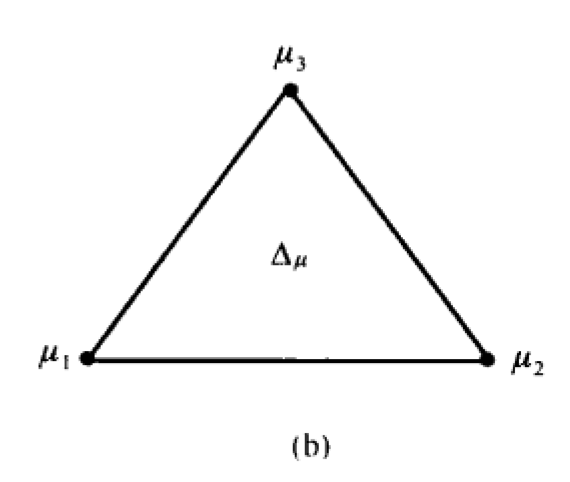
$$a(u,v) = \iint\limits_G \left[\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right]\mathrm{d}x\mathrm{d}y. \tag{6.6.11}$$

第二, 对区域 G 作网格剖分. 常用的剖分有三角剖分和四边形剖分. 这里我们用三角剖分. 首先在 Γ 上取有限个点, 依次联成一闭多边形 Γ_h(h 表示单元的最大直径), 以之近似代替 Γ, 并以 Γ_h 围成的多边形域 G_h 近似代替 G. 选取 Γ_h 时要注意尽量减少几何误差, 必要时需采用曲边元和等参变换法. 然后, 把 G_h 分割成有限个三角单元之和, 从而得到 G_h 的一个三角剖分 (参看图 6.2(a)). 剖分除满足前已提到的一般性要求外, 若有可能从物理直观或理论分析中判断出解 u 变化的剧烈部位和平坦部位, 则在相应部位应使剖分作的 “细密些” 或 “粗糙些”. 究竟应该怎样做剖分才能达到既经济又能保证所需的精度? 这是一个有实际意义又十分困难的问题, 目前主要依靠计算者的实践经验和解的后验估计.

确定网格剖分后, 按一定次序将节点 (单元顶点) 编号. 第 i 号点的坐标记为 (x_i, y_i). 设 $\mu = (\mu_1, \mu_2, \mu_3)$, 用 $\Delta_\mu = \Delta(\mu_1, \mu_2, \mu_3)$ 表示一三角单元, 其顶点的编号为 μ_1, μ_2, μ_3, 并且顶点排列顺序 $\mu_1, \mu_2 \mu_3$ 符合逆时针方向. S_μ 表示 Δ_μ 的面积 (参看图 6.2(b)).



(a)



(b)

图 6.2

第三, 构造基函数或单元形状函数. 假定我们采取线性元, 则每一节点 i 对应一基函数 $\varphi_i.\varphi_i$ 的支集是一切以 i 为顶点的三角单元 $\Delta_\mu = \Delta(i,j,k)$ $(\mu_1 = i, \mu_2 = j, \mu_3 = k)$. 由 §4, 显然 φ_i 在 Δ_μ 的限制为 (参看图 4.1 和 (6.4.3) 式)

$$\varphi_i = L_i = \frac{1}{2S_\mu}(a_i + b_i x + c_i y), \tag{6.6.12}$$

其中

$$a_i = x_j y_k - x_k y_j, \quad b_i = y_j - y_k, \quad c_i = x_k - x_j, \tag{6.6.13}$$

$$2S_\mu = \begin{vmatrix} 1 & x_i & y_i \\ 1 & x_j & y_j \\ 1 & x_k & y_k \end{vmatrix}. \tag{6.6.14}$$

假若与 i 相邻的所有节点是 $l_e (e = 1, 2, \cdots, m)$, 如图 6.3 所示, 则只要在 (6.6.12)~(6.6.14) 中依次取 $j = l_e, k = l_{e+1}$, 其中 $e = 1, 2, \cdots, m, l_{m+1} = l_1$ (图 6.3(a), i 是内点), 或 $e = 1, 2, \cdots, m-1$ (图 6.3(b), i 是界点), 就得出 φ_i 在支集上的表达式.

第四, 形成有限元方程. 假定内点的编号是 $1, 2, \cdots, n_1$, 界点的编号是 $n_1 + 1, n_1 + 2, \cdots, n_1 + n_2$, 则有限元方程为

$$\sum_{i=1}^{n_1} a(\varphi_i, \varphi_j) u_i = (f, \varphi_j) - \sum_{i=n_1+1}^{n_1+n_2} a(\varphi_i, \varphi_j) \alpha_i, \quad j = 1, 2, \cdots, n_1, \tag{6.6.15}$$

其中

$$\alpha_i = \alpha(x_i, y_i), i = n_1 + 1, \cdots, n_1 + n_2,$$

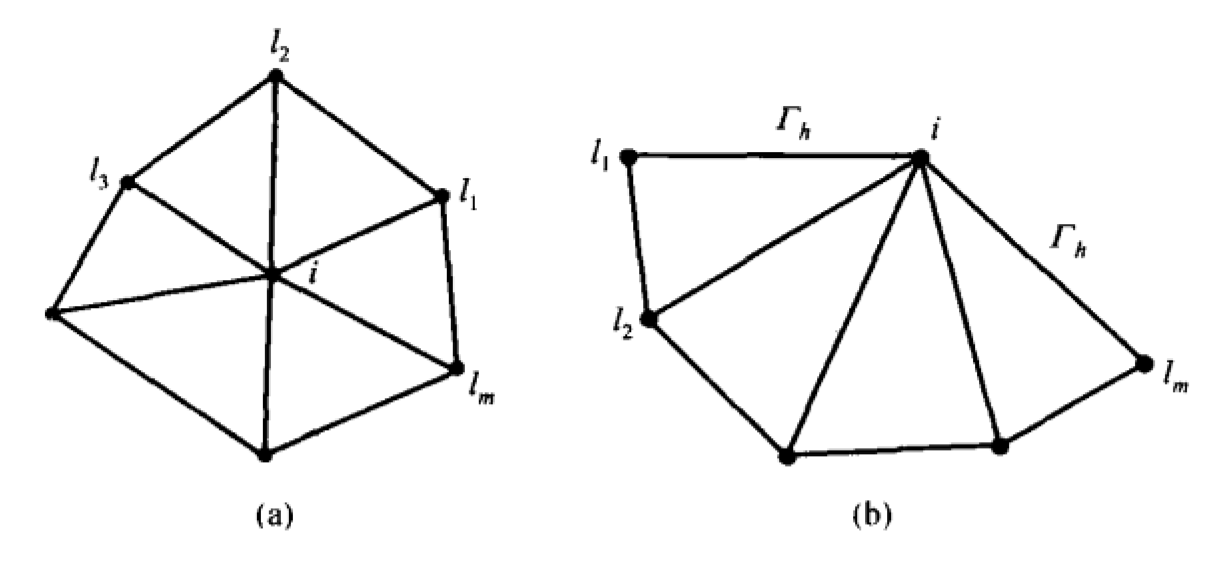


图 6.3

$$(f,\varphi_j)=\iint\limits_G f\varphi_j\mathrm{d}x\mathrm{d}y=\frac{1}{2}\sum_\mu\frac{1}{S_\mu}\iint\limits_{\Delta\mu}(a_j+b_jx+c_jy)f\mathrm{d}x\mathrm{d}y. \tag{6.6.16}$$

$\sum\limits_\mu$ 表示就所有以 j 为顶点的 Δ_μ 求和, 而

$$\begin{aligned} a(\varphi_i,\varphi_j)&=\iint\limits_G\left[\frac{\partial\varphi_i}{\partial x}\cdot\frac{\partial\varphi_j}{\partial x}+\frac{\partial\varphi_i}{\partial y}\cdot\frac{\partial\varphi_j}{\partial y}\right]\mathrm{d}x\mathrm{d}y\\ &=\sum_\mu\frac{1}{4S_\mu^2}\iint\limits_{\Delta\mu}(b_ib_j+c_ic_j)\mathrm{d}x\mathrm{d}y\\ &=\frac{1}{4}\sum_\mu\frac{1}{S_\mu}(b_ib_j+c_ic_j),\\ &\quad i=1,2,\cdots,n_1+n_2,\quad j=1,2,\cdots,n_1. \end{aligned} \tag{6.6.17}$$

这里 $\sum\limits_\mu$ 表示就所有以 $\overline{ij}$ 为边的三角单元求和 ($i=j$ 时是以 j 为顶点的单元求和). 出现在 (6.6.16), (6.6.17) 中的 a_i,b_i,c_i 及 S_μ 均按公式 (6.6.13), (6.6.14) 计算.

第五, 有限元方程的解法. 根据问题的规模和计算机的容量、速度, 选择适当的解法, 如消元法, 第二章介绍的 SOR 法, 共轭梯度法以及本章 §7 将要介绍的多重网格法.

第六, 编制程序, 上机计算.

表 6.1 求积公式系数及节点表

W_i	L_1	L_2	L_3	重数
	3 点公式	2 阶精度		
0.333 333 333 333 333	0.666 666 666 666 667	0.166 666 666 666 667	0.166 666 666 666 667	3
	3 点公式	2 阶精度		
0.333 333 333 333 333	0.500 000 000 000 000	0.500 000 000 000 000	0.000 000 000 000 000	3
	4 点公式	3 阶精度		
−0.562 500 000 000 000	0.333 333 333 333 333	0.333 333 333 333 333	0.333 333 333 333 333	1
0.520 833 333 333 333	0.600 000 000 000 000	0.200 000 000 000 000	0.200 000 000 000 000	3
	6 点公式	3 阶精度		
0.166 666 666 666 667	0.659 027 622 374 092	0.231 933 368 553 031	0.109 039 009 072 877	6
	6 点公式	4 阶精度		
0.109 951 743 655 322	0.816 847 572 980 459	0.091 576 213 509 771	0.091 576 213 509 771	3
0.223 381 589 678 011	0.108 103 018 168 070	0.445 948 490 915 965	0.445 948 490 915 965	3
	7 点公式	4 阶精度		
0.375 000 000 000 000	0.333 333 333 333 333	0.333 333 333 333 333	0.333 333 333 333 333	1
0.104 166 666 666 667	0.736 712 498 968 435	0.237 932 366 472 434	0.025 355 134 551 932	6
	7 点公式	5 阶精度		
0.225 030 000 330 000	0.333 333 333 333 333	0.333 333 333 333 333	0.333 333 333 333 333	1
0.125 939 180 544 827	0.797 426 985 353 087	0.101 286 507 323 456	0.101 286 507 323 456	3
0.132 394 152 788 506	0.470 142 064 105 115	0.470 142 064 105 115	0.059 715 871 789 770	3
	9 点公式	5 阶精度		
0.205 950 504 760 887	0.124 649 503 233 232	0.437 525 248 383 384	0.437 525 248 383 384	3
0.063 691 414 286 223	0.797 112 651 860 071	0.165 409 927 389 841	0.037 477 420 750 088	6
	12 点公式	6 阶精度		
0.050 844 906 370 207	0.873 821 971 016 996	0.063 089 011 491 502	0.063 089 011 491 502	3
0.116 786 275 726 379	0.501 426 509 658 179	0.249 286 745 170 910	0.249 286 745 170 911	3
0.082 851 075 618 374	0.636 502 499 121 399	0.310 352 451 033 785	0.053 145 049 844 816	6
	13 点公式	7 阶精度		
−0.149 570 044 467 670	0.333 333 333 333 333	0.333 333 333 333 333	0.333 333 333 333 333	1
0.175 615 257 433 204	0.479 308 067 841 923	0.260 345 966 079 038	0.260 345 966 079 038	3
0.053 347 235 608 839	0.869 739 794 195 598	0.065 130 102 902 216	0.065 130 102 902 216	3
0.077 113 760 890 257	0.638 444 188 569 809	0.312 865 496 004 875	0.486 903 154 253 160	6

6.5 数值例子

求解 Helmholtz 方程:

$$\begin{cases} \Delta u + k^2 u = -1, & \text{于 } G, \\ u = 0, & \text{于 } \Gamma_1, \\ \dfrac{\partial u}{\partial \boldsymbol{\nu}} = 0, & \text{于 } \Gamma_2, \end{cases}$$

其中 $k=5, G=(0,1)\times(0,1)$, $\Gamma_1 = \overline{AC}\bigcup\overline{CD}, \Gamma_2 = \overline{AB}\bigcup\overline{BD}, \boldsymbol{\nu}$ 是 Γ_2 的单位外

法向. (参看图 6.4).

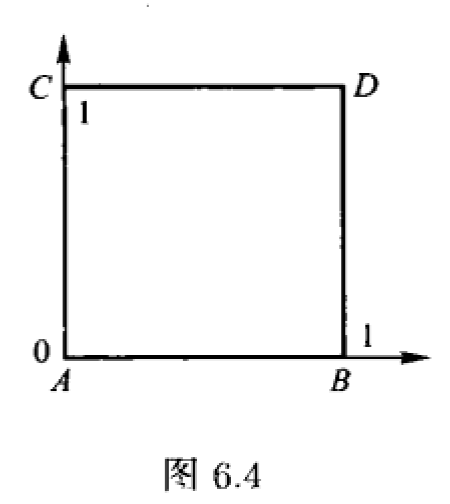


图 6.4

我们用线性有限元法 (FEM), 有限体积法 (FVM) 和谱方法求解. 取网格步长 $h = 1/N$, 先作正方形网, 再作与东西向平行的对角线, 将它剖分成三角形网. 就 $N = 64, 128, 256$ 用 FEM, FVM 和谱方法计算, 计算结果如表 6.2. 因为解关于对角线 $\overline{BC}: y = 1 - x$ 对称, 表 6.2 只列出 $\overline{BC}$ 上方的部分数据. 从中看出, 谱方法收敛最快, 精确到小数后四位,FEM 和 FVM 的精度接近, 分别由上方和下方逼近谱方法的解.

表 6.2

(x, y)	N	FEM	FVM	谱方法
(1/8,7/8)	64	−0.365 136	−0.339 712	−0.350 740
	128	0.354 236	−0.347 920	−0.350 751
	256	−0.351 609	−0.350 032	−0.350 737
(1/8,31/32)	64	−0.093 005	−0.086 475	−0.089 329
	128	−0.090 213	−0.088 591	−0.089 311
	256	−0.089 540	−0.089 135	−0.089 308
(1/4,3/4)	64	−1.214 295	−1.130 888	1.166 859
	128	−1.178 386	−1.157 668	−1.166 878
	256	1.169 731	−1.164 559	−1.166 876
(1/4,15/16)	64	−0.341 764	−0.318 056	−0.328 286
	128	0.331 582	0.325 693	−0.328 328
	256	−0.329 129	−0.327 658	−0.328 320
(1/2,1/2)	64	−1.772 112	−1.653 661	−1.704 545
	128	−1.720 947	−1.691 527	−1.794 545
	256	−1.708 616	−1.701 271	−1.704 545
(1/2,3/4)	64	−1.607 561	−1.498 246	−1.545 257
	128	−1.560 386	−1.533 235	−1.545 265
	256	−1.549 017	−1.542 239	−1.545 264

续表

(x,y)	N	FEM	FVM	谱方法
(1/2,7/8)	64	−0.916 827	−0.854 048	−0.881 063
	128	−0.889 745	−0.874 152	−0.881 068
	256	−0.883 218	−0.879 326	−0.881 061
(1,0)	64	3.384 657	3.152 834	3.254 590
	128	3.285 971	3.228 553	3.254 586
	256	3.262 327	3.248 038	3.254 584
(1,1/2)	64	−0.062 681	−0.063 598	−0.063 599
	128	−0.063 379	−0.063 602	−0.063 601
	256	−0.063 547	−0.063 602	−0.063 602
(1,3/4)	64	−0.962 919	−0.899 006	−0.926 543
	128	−0.935 381	−0.919 506	−0.926 554
	256	−0.928 745	−0.924 781	−0.926 554
(1,7/8)	64	−0.632 165	−0.589 335	−0.607 682
	128	−0.613 629	−0.602 992	−0.607 690
	256	−0.609 162	−0.606 506	−0.607 684

习　　题

(实习题) 用线性元求下列问题的数值解:

$$\begin{cases} \Delta u = -2, \ -1 < x, \ y < 1, \\ u(x,-1) = u(x,1) = 0, \ -1 < x < 1, \\ u_x(-1,y) = 1, u_x(1,y) = 0, \ -1 < y < 1 \end{cases}$$

(精确到小数点后第 6 位).

*§7　多重网格法

多重网格法是一种解差分方程, 有限元方程的高效率的算法. 多重网格法的网格剖分不是一次取定, 而是由粗到细取一系列网格, 计算时在粗网格求初值 (粗网格校正), 在细网格进行迭代 (光滑化). 我们仅以模型问题为例介绍方法的计算框架.

7.1　差分形式的二重网格法

考虑模型问题:

$$-u''(x) = f(x), \quad x \in G = (0,1), \tag{6.7.1}$$
$$u(x) = 0, \quad x \in \Gamma = \{0,1\}.$$

取步长

$$h_j = 2^{-j-1}(j = 0,1,\cdots,J)$$

和节点

$$G_j = \{x_i = ih_j, i = 1,2,\cdots,N_j = 2^{j+1}-1\}, \quad j = 0,1,\cdots,J.$$

关于 G_J 构造逼近 (6.7.1) 的差分格式:

$$L_J u_i = -(u_{i+1} - 2u_i + u_{i-1})/h_J^2 = f_i \quad (f_i = f(x_i)), \tag{6.7.2}$$
$$u_0 = u_{N_J+1} = 0, \quad i = 1,2,\cdots,N_J,$$

记 $\boldsymbol{U}_J = (u_1,\cdots,u_{N_J})^{\mathrm{T}}, \boldsymbol{F}_J = (f_1,\cdots,f_{N_J})^{\mathrm{T}}, \boldsymbol{A}_J$ 是 (6.7.2) 的系数矩阵, 将 (6.7.2) 写成矩阵形式:

$$\boldsymbol{A}_J\boldsymbol{U}_J = \boldsymbol{F}_J \text{ 或 } \frac{1}{2}h_J^2(\boldsymbol{A}_J\boldsymbol{U}_J - \boldsymbol{F}_J) = 0. \tag{6.7.3}$$

引进松弛因子 2ω, 构造迭代

$$\boldsymbol{U}_J^{(m+1)} = \boldsymbol{U}_J^{(m)} - \omega h_J^2(\boldsymbol{A}_J\boldsymbol{U}_J - \boldsymbol{F}_J). \tag{6.7.4}$$

由第二章 §6 例 6.1, 知 $\boldsymbol{A}_J$ 的特征向量为

$$\boldsymbol{e}_\mu = (\sin \nu\mu h_J)_{\nu=1}^{N_J}, \quad \mu = 1,2,\cdots,N_J, \tag{6.7.5}$$

相应的特征值是

$$4h_J^{-2}\sin^2\frac{1}{2}\mu\pi h_J, \quad \mu = 1,2,\cdots,N_J.$$

迭代矩阵 $\boldsymbol{M}_J = \boldsymbol{I} - \omega h_J^2\boldsymbol{A}_J$ 的特征向量仍是 (6.7.5), 特征值为

$$\lambda_\mu(\omega) = 1 - 4\omega\sin^2\frac{1}{2}\mu\pi h_J. \tag{6.7.6}$$

m 次迭代误差

$$\boldsymbol{U}_J^{(m)} - \boldsymbol{U}_J = \sum_{\mu=1}^{N_J} c_\mu\lambda_\mu^m(\omega)\boldsymbol{e}_\mu, \quad m = 1,2,\cdots. \tag{6.7.7}$$

显然迭代 (6.7.4) 当 $\omega \in \left(0, \dfrac{1}{2}\right)$ 时 $|\lambda_\mu| < 1$, 从而 $\boldsymbol{U}_J^{(m)} - \boldsymbol{U}_J \to 0 (m \to \infty)$. 现考虑 (6.7.7) 右端各项趋于 0 的速度. 在 (6.7.7) 右端的和式中, 大 μ 对应的是高频项 (参看 (6.7.5)), 相应的 $|\lambda_\mu| \leqslant \rho < 1$, 例如当 $\mu h_J \geqslant \dfrac{1}{2}\left(\mu \geqslant \dfrac{N_J+1}{2}\right)$ 时, $\left|\sin \dfrac{1}{2}\mu\pi h_J\right| \geqslant \dfrac{1}{2}$, 从而

$$|\lambda_\mu(\omega)| \leqslant |1-2\omega| < 1, \quad \omega \in \left(0, \frac{1}{2}\right).$$

特别 $\left|\lambda_\mu\left(\dfrac{1}{4}\right)\right| \leqslant \dfrac{1}{2}$, 故在迭代过程中高频项很快衰减到零. 反之, (6.7.7) 中小 μ 对应的是低频项, 相应的 $|\lambda_\mu|$ 接近 1 (当 $h_j \to 0$), 例如恒有

$$\lambda_1(\omega) \approx 1 - O(h_J^2), \quad h_J \to 0 (J \to \infty).$$

故在迭代过程中低频项衰减的速度很慢, 这就是各种迭代法的敛速随网格变细而减小的根源 (参看第二章 §6).

另一方面, 从 (6.7.7) 还知道, 随着高频项衰减, 迭代误差 $\boldsymbol{E}_J = \boldsymbol{U}_J^{(m)} - \boldsymbol{U}_J$ 将变光滑. 令

$$\boldsymbol{D}_J = \boldsymbol{A}_J(\boldsymbol{U}_J^{(m)} - \boldsymbol{U}_J) = \boldsymbol{A}_J\boldsymbol{U}_J^{(m)} - \boldsymbol{F}_J, \tag{6.7.8}$$

它是 m 次迭代的残量, 则 $\boldsymbol{E}_J$ 满足

$$\boldsymbol{A}_J\boldsymbol{E}_J = \boldsymbol{D}_J. \tag{6.7.9}$$

这是和 (6.7.3) 相同的差分方程, 只不过右端换成残量 $\boldsymbol{D}_J$. 因为 $\boldsymbol{E}_J$ 光滑, 它比 $\boldsymbol{U}_J$ 更易逼近, 所以可在粗网格 G_{J-1} 上求解 $\boldsymbol{E}_J$. 为此, 引进由 G_J 上的网格函数到 G_{J-1} 上的网格函数的限制算子 r:

$$\boldsymbol{D}_{J-1} = r\boldsymbol{D}_J, \tag{6.7.10}$$

并用粗网格 G_{J-1} 上的方程

$$\boldsymbol{A}_{J-1}\boldsymbol{E}_{J-1} = \boldsymbol{D}_{J-1} \tag{6.7.11}$$

代替 (6.7.9). 求出 $\boldsymbol{E}_{J-1} = \boldsymbol{A}_{J-1}^{-1}\boldsymbol{D}_{J-1}$ 后, 再利用由粗网格 G_{J-1} 到细网格 G_J 的延拓算子 p, 得到

$$\boldsymbol{E}_J = p\boldsymbol{E}_{J-1}. \tag{6.7.12}$$

从而得到 $\boldsymbol{U}_J$ 的一次校正:

$$\boldsymbol{U}_J = \boldsymbol{U}_J^{(m)} - \boldsymbol{E}_J. \tag{6.7.13}$$

将此校正值作为 G_J 上的初始近似再光滑化 (一般用迭代法), 就完成了二重网格的一次迭代. 如此反复迭代, 直至收敛为止, 这就是二重网格法.

关于限制算子 r, 最平凡的取法是

$$\boldsymbol{D}_{J-1}(x)=(r\boldsymbol{D}_J)(x)=\boldsymbol{D}_J(x),\quad x\in G_{J-1}\subset G_J.$$

另一种是取邻点的权平均, 例如

$$(r\boldsymbol{D}_J)(x)=\frac{1}{4}[\boldsymbol{D}_J(x-h_J)+2\boldsymbol{D}_J(x)+\boldsymbol{D}_J(x+h_J)],\quad x\in G_{J-1}.$$

关于延拓算子 p, 最简单的取法是分片线性插值, 即

$$(p\boldsymbol{E}_{J-1}(x))=\begin{cases}\boldsymbol{E}_{J-1}(x),\ x\in G_{J-1},\\ \dfrac{1}{2}[\boldsymbol{E}_{J-1}(x-h_J)+\boldsymbol{E}_{J-1}(x+h_J)],\quad x\in G_J-G_{J-1},\end{cases}$$

$$\boldsymbol{E}_{J-1}(0)=\boldsymbol{E}_{J-1}(1)=0.$$

综合上述, 实现二重网格法的主要步骤如下:

(1) 光滑化. 以 $\boldsymbol{U}_J^{(m)}$ 为初始近似, 用 (6.7.4) 迭代若干次得 $\overline{\boldsymbol{U}}$, 例如迭代一次, 则

$$\overline{\boldsymbol{U}_J}=\boldsymbol{U}_J^{(m)}-\omega h_J^2(\boldsymbol{A}_J\boldsymbol{U}_J^{(m)}-\boldsymbol{F}_J). \tag{6.7.14}$$

(2) 粗网格校正.

$$\boldsymbol{D}_J=\boldsymbol{A}_J\overline{\boldsymbol{U}_J}-\boldsymbol{F}_J\quad (\text{计算残量}), \tag{6.7.15$_1$}$$

$$\boldsymbol{D}_{J-1}=r\boldsymbol{D}_J\quad (\text{限制残量}), \tag{6.7.15$_2$}$$

$$\boldsymbol{E}_{J-1}=\boldsymbol{A}_{J-1}^{-1}\boldsymbol{D}_{J-1}\quad (\text{粗网格计算}), \tag{6.7.15$_3$}$$

$$\overline{\overline{\boldsymbol{U}_J}}=\overline{\boldsymbol{U}_J}-p\boldsymbol{E}_{J-1}\quad (\text{校正}), \tag{6.7.15$_4$}$$

可缩写为

$$\overline{\overline{\boldsymbol{U}_J}}=\overline{\boldsymbol{U}_J}-p\boldsymbol{A}_{J-1}^{-1}r(\boldsymbol{A}_J\overline{\boldsymbol{U}_J}-\boldsymbol{F}_J). \tag{6.7.15}$$

(3) 光滑化. 以 $\overline{\overline{\boldsymbol{U}_J}}$ 代替 (6.7.14) 的 $\boldsymbol{U}_J^{(m)}$, 得

$$\boldsymbol{U}_J^{(m+1)}=\overline{\overline{\boldsymbol{U}_J}}-\omega h_J^2(\boldsymbol{A}_J\overline{\overline{\boldsymbol{U}_J}}-\boldsymbol{F}_J). \tag{6.7.16}$$

如此循环迭代, 直至在要求精度内收敛为止. 迭代是在相邻网格 G_{J-1}, G_J 之间进行, 故称为二重网格法.

7.2 有限元形式的二重网格法

有限元形式与差分形式的二重网格法没有本质区别, 但生成限制和延拓算子更为简便. 对问题 (6.7.1) 引进双线性形式:

$$a(u,v)=\int_0^1 u'(x)v'(x)\mathrm{d}x.$$

则变分形式为: 求 $u\in H_0^1(G)$, 满足

$$a(u,v)=(f,v),\quad \forall v\in H_0^1(G). \tag{6.7.17}$$

如前取步长 h_j 和剖分节点 $G_j=\{x_i=ih_j: i=1,2,\cdots,N_j=2^{j+1}-1\}, j=1,2,\cdots,J$. 第 k 层剖分的单元 $K_i^{(k)}=[x_{i-1},x_i], x_i=ih_k$. 设 $M_k\subset H_0^1(G)$ 是剖分 $\{K_i^{(k)}\}$ 上的分片线性元空间, 显然 $M_1\subset M_2\subset\cdots\subset M_J$. 以 M_J 为有限元空间, 求解 (6.7.17) 的有限元法为: 求 $u_J\in M_J$ 使

$$a(u_J,\phi_i)=(f,\phi_i),\quad i=1,2,\cdots,N_J, \tag{6.7.18}$$

其中 $\phi_i\in M_J, \phi_i(x_j)=\delta_{ij}, x_j=jh_J, u_J=\sum\limits_{j=1}^{N_J}u_j\phi_j$. 容易验证, (6.7.18) 就是差分格式 (6.7.2) (仅右端略有不同). 设算子 $A_k: M_k\to M_k$ 由下式定义:

$$(A_ku,v)=a(u,v),\quad \forall u,v\in M_k,\quad k=1,2,\cdots,J. \tag{6.7.19}$$

则 (6.7.18) 就是

$$A_Ju_J=f_J,\quad f_J=\sum_{i=1}^{N_J}(f,\phi_i)\phi_i. \tag{6.7.20}$$

现在有了一系列子空间

$$M_1\subset M_2\subset\cdots\subset M_J$$

和算子 $A_k: M_k\to M_k, k=1,2,\cdots,J, A_k$ 于 M_k 对称正定. 由 (6.7.19) 还知道 $(A_ku,v)=(A_{k+1}u,v)$, 当 $u,v\in M_k$. 设 Q_k 是 M_{k+1} 往 M_k 的正投影: $(Q_ku,v)=(u,v), v\in M_k$, 则 $Q_kA_{k+1}u=A_ku$. 显然 A_J 与差分算子 L_J (参看 (6.7.2)) 的特征向量相同, 其特征值 $\Lambda_\mu^{(J)}$ 则是 L_J 的 h_J 倍, 即

$$\Lambda_\mu^{(J)}=4h_J^{-1}\sin^2\frac{1}{2}\mu\pi h_j,\quad \mu=1,2,\cdots,N_J. \tag{6.7.21}$$

当 $h_J \to 0$ 时 $\Lambda_1^{(J)} \approx h_J, \Lambda_{N_J}^J \approx O(h_J^{-1})$, 从而 $\mathrm{Cond}(A_J) \approx h_J^{-2}$. 仿照 7.1 节, 可构造第 J 层和第 $J-1$ 层之间的二重网格迭代. 现在限制算子可取为 M_J 到 M_{J-1} 的正投影 Q_{J-1}:

$$(Q_{J-1}u, v) = (u, v), \quad \forall v \in M_{J-1}. \tag{6.7.22}$$

延拓算子的一种自然取法是 $p = I : M_{J-1} \to M_{J-1} \subset M_J$. 于是二重网格迭代为:

(1) 光滑化. 对给定的初始近似 $u_J^{(m)}$, 作迭代

$$\overline{u}_J = u_J^{(m)} - \omega h_J (A_J u_J^{(m)} - f_J). \tag{6.7.23}$$

(2) 粗网格校正:

$$(6.7.24)_1 \qquad D_J = A_J \overline{u}_J - f_J \quad (计算残量),$$
$$(6.7.24)_2 \qquad D_{J-1} = Q_{J-1} D_J \quad (限制残量),$$
$$(6.7.24)_3 \qquad E_{J-1} = A_{J-1}^{-1} D_{J-1} \quad (粗网格计算),$$
$$(6.7.24)_4 \qquad \overline{\overline{u}}_J = \overline{u}_J - E_{J-1} \quad (校正),$$

可缩写为

$$\overline{\overline{u}}_J = \overline{u}_J - A_{J-1}^{-1} Q_{J-1} (A_J \overline{u}_J - f_J). \tag{6.7.24}$$

(3) 光滑化

$$u_J^{(m+1)} = \overline{\overline{u}}_J - \omega h_J (A_J \overline{\overline{u}}_J - f_J). \tag{6.7.25}$$

如此循环迭代, 直至在要求精度内收敛为止.

7.3 多重网格迭代和套迭代技术

二重网格法是收敛很快的算法, 但在粗网格校正步 $(6.7.15)_3$ 或 $(6.7.24)_3$ 中, 需求下列方程的解:

$$A_{J-1} E_{J-1} = D_{J-1}. \tag{6.7.26}$$

这对高维边值问题可能仍不经济, 为此可将二重网格迭代用于 $J-1$ 和 $J-2$ 层, 得到包括 $J, J-1$ 和 $J-2$ 层的三重网格法. 假定在 $J-1, J-2$ 层用 γ 次二重网格迭代解 (6.7.26), 则在 $J-2$ 层需 γ 次解方程 $A_{J-2}E_{J-2} = D_{J-2}$(γ 次修改 E_{J-2}). 为此又需在 $J-2, J-3$ 层用二重网格法. 如此下去, 直至包括 $J+1$ 层 $\{J, J-1, \cdots, 1, 0\}$. 第 0 层对应初始网格 G_0, 可精确求解 $\boldsymbol{A}_0\boldsymbol{E}_0 = \boldsymbol{D}_0$ (也可用

迭代法). 如此的 $(J+1)$ 层迭代称为多重网格法 (Multi-Grid Method), 简记为 MGM.

以上我们是用 $J-1$ 层的 γ 次多重网格迭代代替精确求解 (6.7.26). 在 J 层调用一次 MGM, 会引起在 $J-1$ 层调用 γ 次 MGM, 进而引起在 $J-2$ 层调用 γ^2 次 MGM, 如此等等. 通常取 $\gamma=1$ 或 2. 取 $J=4, \gamma=1$ 时的运算顺序如图 7.1. 五线谱符号表示层数, 其形状似 V, 故也称为 V 循环. $\gamma=2$ 的运算顺序如图 7.2, 其形状似 W, 故也称为 W 循环.

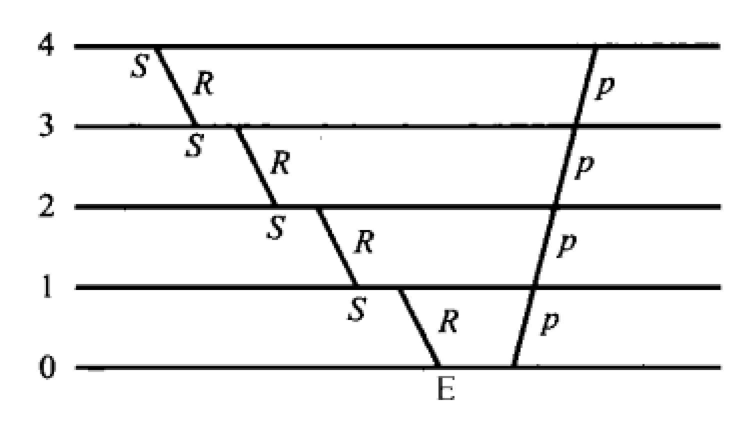


图 7.1 $\gamma=1, J=4$ 的一次多重网格迭代

S: 光滑步; R: 残量限制; E: 0 层求解; p: 校正

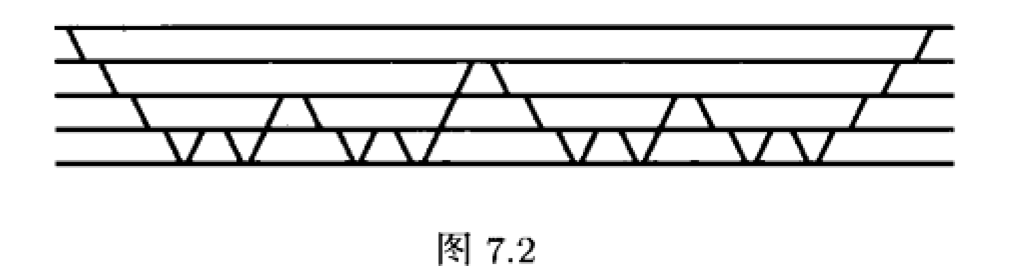


图 7.2

多重网格法是一个由细网到粗网的计算过程, 为在细网格 G_J 上解 $\boldsymbol{A}_J\boldsymbol{U}_J=\boldsymbol{F}_J$, 把它化为在一系列粗网格 $G_{J-1}, G_{J-2}, \cdots, G_0$ 上求解. 与此相反, 数值分析常用的一种所谓套迭代法是一个由粗网到细网的计算过程, 粗网的计算结果为下一层细网的计算提供初始近似. 套迭代和任一种迭代法结合 (比如 SOR 法) 就产生各种套迭代技术. 由套迭代法逐步得到的近似解 $\overline{u}_0, \overline{u}_1, \cdots, \overline{u}_{J-1}$ 可用于估计误差, 为自动选步长提供依据; 还可用外推技术提高解的精度 (参看 [18],[19]).

§8 初边值问题的有限元法

Galerkin 法以及由此发展起来的有限元法也可用于解初边值问题 (非驻定问题), 包括抛物型方程和双曲方程. 此时将时间变量 t 看成参数, 用虚功原理将初边值问题 化成变分形式, 然后用 Galerkin 有限元法求解.

8.1 热传导方程

考虑热传导方程的初边值问题:

$$(6.8.1)_1 \qquad \frac{\partial u}{\partial t} = \Delta u + f, (x, y) \in G \subset \mathbf{R}^2, t > 0,$$

$$(6.8.1)_2 \qquad u(x, y; 0) = \psi(x, y), (x, y) \in G,$$

$$(6.8.1)_3 \qquad u|_{\partial G} = 0, \quad t > 0,$$

其中 $\Delta u = \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2}, f = f(x, y), G$ 是具光滑边界 ∂G 的平面有界域. 设对固定的 $t > 0$, 解 $u(x, y; t)$ 关于 (x, y) 属于 $C^2(\overline{G})$. 以 $v \in H_0^1(G)$ 乘 $(6.8.1)_1$ 两端并积分, 得

$$(6.8.2) \qquad \int_G \left(\frac{\partial u}{\partial t} - \Delta u - f\right) v \mathrm{d}x \mathrm{d}y = 0.$$

利用 Green 公式和边值条件 $(6.8.1)_3$, 得

$$(6.8.3) \qquad \int_G \left(\frac{\partial u}{\partial t} v + \frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} - fv\right) \mathrm{d}x \mathrm{d}y = 0.$$

引进双线性形式和 $L^2(G)$ 内积:

$$a(u, v) = \int_G \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right) \mathrm{d}x \mathrm{d}y,$$

$$(f, v) = \int_G f v \mathrm{d}x \mathrm{d}y.$$

则初边问题 $(6.8.1)_1$–$(6.8.1)_3$ 的变分形式为: 求 $u(x, y : t) \in H_0^1(G)$ (视 t 为参数), 满足

$$(6.8.4)_1 \qquad \left(\frac{\partial u}{\partial t}, v\right) + a(u, v) = (f, v), \quad \forall v \in H_0^1(G),$$

$$(6.8.4)_2 \qquad u(x, y; 0) = \psi(x, y).$$

并称如此的 u 为初边问题 (6.8.1) 的广义解.

在 H_0^1 中取一 n 维子空间 U_h, 所谓 Galerkin 法就是求含参数 t 的函数 $u_h(x, y; t) \in U_h (t \geqslant 0)$, 满足

$$(6.8.5)_1 \qquad \left(\frac{\partial u_h}{\partial t}, v_h\right) + a(u_h, v_h) = (f, v_h), \quad \forall v_h \in U_h,$$

$$(6.8.5)_2 \qquad (u(x, y; 0), v_h) = (\psi, v_h), \quad \forall v_h \in U_h.$$

在 U_h 中取定一组基底 $\phi_1, \phi_2, \cdots, \phi_n$, 将 u_h 表为

$$u_h(x, y; t) = \sum_{i=1}^{n} \mu_i(t) \phi_i(x, y),$$

代到 $(6.8.5)_1$, 并取 $v_h = \phi_j$, 就得到关于 $\mu_1(t), \cdots, \mu_n(t)$ 的常微分方程组:

$$\sum_{i=1}^{n}(\phi_i, \phi_j)\frac{\mathrm{d}\mu_i}{\mathrm{d}t} + \sum_{i=1}^{n} a(\phi_i, \phi_j)\mu_i = (f, \phi_j), \quad j = 1, 2, \cdots, n. \tag{6.8.6}$$

初始条件可按 $(6.8.5)_2$ 或其他方法给出. 若按 $(6.8.5)_2$ 取初值 $u_h(x, y; 0) = \sum_{i=1}^{n}\mu_i(0)\phi_i$, 则 $\mu_i(0)$ 由下列方程组确定:

$$\sum_{i=1}^{n}(\phi_i, \phi_j)\mu_i(0) = (\psi, \phi_j), \quad j = 1, 2, \cdots, n. \tag{6.8.7}$$

至此我们得到了**半离散化 Galerkin 方程** $(6.8.5)_1$ 或 (6.8.6). 若进一步对时间 t 离散化, 就得到**全离散化 Galerkin 法**. 例如用向前差商 (向前 Euler 格式),

$$\frac{1}{\tau}(u_h^{n+1} - u_h^n, v_h) + a(u_h^n, v_h) = (f, v_h), \quad n = 0, 1, \cdots, \tag{6.8.8}$$

用向后差商 (向后 Euler 格式),

$$\frac{1}{\tau}(u_h^{n+1} - u_h^n, v_h) + a(u_h^{n+1}, v_h) = (f, v_h), \quad n = 0, 1, \cdots, \tag{6.8.9}$$

或用 Crank-Nicolson 格式 (改进的 Euler 折线法)

$$\frac{1}{\tau}(u_h^{n+1} - u_h^n, v_h) + \frac{1}{2}a(u_h^{n+1} + u_h^n, v_h) = (f, v_h), \quad n = 0, 1, \cdots, \tag{6.8.10}$$

其中上标 n 表示在 $t = t_n = n\tau$ 的近似.

特别若对域 G 作三角剖分, 并取 $U_h \subset H_0^1$ 为分片多项式函数空间, 则前述 Galerkin 法就是有限元法. 以前各节关于有限元空间的构造和算法都可用到抛物方程.

注意全离散格式是由常微分方程组 $(6.8.5)_1$ 或 (6.8.6) 离散化得到的, 所以需要讨论它的稳定性 (参看第一章). 作为例子, 我们证明格式 (6.8.10) 关于初值稳定. 在 (6.8.10) 中取 $v_h = u_h^{n+1} - u_h^n, f = 0$, 得

$$\frac{1}{\tau}\|u_h^{n+1} - u_h^n\|^2 + \frac{1}{2}a(u_h^{n+1} + u_h^n, u_h^{n+1} - u_h^n) = 0, \tag{6.8.11}$$

其中 $\|\cdot\|$ 表示 L^2 范数. 利用 $a(\cdot,\cdot)$ 的对称性得

$$a(u_h^{n+1} + u_h^n, u_h^{n+1} - u_h^n) = a(u_h^{n+1}, u_h^{n+1}) - a(u_h^n, u_h^n) = |u_h^{n+1}|_1^2 - |u_h^n|_1^2.$$

其中 $|\ \ |_1$ 是半模. 于是由 (6.8.11) 得

$$|u_h^{n+1}|_1 \leqslant |u_h^n|_1 \leqslant \cdots \leqslant |u_h^0|_1,$$

由 Poincaré 不等式知稳定性得证.

8.2 波动方程

考虑波动方程的初边值问题:

$$(6.8.12)_1 \qquad \frac{\partial^2 u}{\partial t^2} = \Delta u, (x,y) \in G \subset \mathbf{R}^2, \quad t > 0,$$

$$(6.8.12)_2 \qquad u(x,y;0) = g_0(x,y), \quad u_t(x,y;0) = g_1(x,y),$$

$$(6.8.12)_3 \qquad u|_{\partial G} = 0, \quad t > 0,$$

其中 $\Delta u = \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2}$, G 是具光滑边界 ∂G 的平面有界域. 与前述方法类似可得到 $(6.8.12)_{1-2-3}$ 的变分形式

$$(6.8.13)_1 \qquad \left(\frac{\partial^2 u}{\partial t^2}, v\right) + a(u,v) = 0, \forall v \in H_0^1(G),$$

$$(6.8.13)_2 \qquad u(x,y;0) = g_0(x,y), u_t(x,y;0) = g_1(x,y),$$

并称如此的解 u 为初边值问题 $(6.8.12)_{1-2-3}$ 的广义解.

在 H_0^1 中取一 n 维子空间 U_h. 所谓 Galerkin 法就是求含参数 t 的函数 $u_h(x,y;t) \in U_h (t \geqslant 0)$, 满足

$$(6.8.14)_1 \qquad \left(\frac{\partial^2 u_h}{\partial t^2}, v_h\right) + a(u_h, v_h) = 0, \quad \forall v_h \in U_h,$$

$$(6.8.14)_2 \qquad (u_h(x,y;0), v_h) = (g_0, v_h), \left(\frac{\partial u_h}{\partial t}(x,y;0), v_h\right) = (g_1, v_h), \quad \forall v_h \in U_h.$$

在 U_h 中取定一组基底 $\varphi_1, \varphi_2, \cdots, \varphi_n$ 将 u_h 表为

$$u_h(x,y,t) = \sum_{i=1}^n \mu_i(t) \varphi_i(x,y),$$

代到 $(6.8.14)_{1-2}$, 并取 $v_h = \varphi_j$, 就得到关于 $\mu_1, \cdots, \mu_n$ 的常微分方程组:

$$(6.8.15)_1 \qquad \sum_{i=1}^n (\varphi_i, \varphi_j) \frac{\mathrm{d}^2 \mu_i}{\mathrm{d}t^2} + \sum_{i=1}^n a(\varphi_i, \varphi_j) \mu_i = 0, \quad j = 1, 2, \cdots, n.$$

初始条件为

$$(6.8.15)_2 \qquad \mu_j(0) = \alpha_j, \quad \frac{\partial \mu_j(0)}{\partial t} = \beta_j.$$

其中 $\mu_j(0)$ 和 $\dfrac{\partial \mu_j(0)}{\partial t}$ 按 $(6.8.14)_{1-2}$ 确定:

$$(6.8.16) \qquad \sum_{i=1}^n (\varphi_i, \varphi_j) \mu_i(0) = (g_0, \varphi_j),$$

$$\sum_{i=1}^n (\varphi_i, \varphi_j) \frac{\partial \mu_i(0)}{\partial t} = (g_1, \varphi_j), \quad j = 1, 2, \cdots, n.$$

这样我们就得到**半离散化 Galerkin 方程** $(6.8.15)_1$–(6.8.16). 这是一个二阶常微分方程组.

为得到全离散化 Galerkin 方程, 需用差分法对时间 t 进一步离散. 设时间步长为

$$\tau > 0, \quad t_n = n\tau(n = 0, 1, \cdots, N; N\tau = T),$$

$u_h^n = u_h(x, y; t_n)$. 引进差商符号

$$\begin{aligned} \partial_{t\bar{t}} u_h^n &= (u_h^{n+1} - 2u_h^n + u_h^{n-1})/\tau^2, \\ u_h^{n,\frac{1}{4}} &= (u_h^{n+1} + 2u_h^n + u_h^{n-1})/4, \end{aligned}$$

则得**全离散 Galerkin 方程**, 例如

显格式

$$(\partial_{t\bar{t}} u_h^n, v_h) + a(u_h^n, v_h) = 0, \quad \forall v_h \in U_h, \tag{6.8.17}$$

隐格式

$$(\partial_{t\bar{t}} u_h^n, v_h) + a\left(u_h^{n,\frac{1}{4}}, v_h\right) = 0, \quad \forall v_h \in U_h. \tag{6.8.18}$$

特别若取 U_h 为有限元空间, 则得 Galerkin 有限元法. 可以证明显式 (6.8.17) 条件稳定, 隐式 (6.8.18) 恒稳定 (参看 [8]).



名 词 索 引

R

S

T

V

W

X



参考文献

[1] 伍卓群, 李勇. 常微分方程. 北京: 高等教育出版社, 2004.

[2] 江泽坚, 孙善利. 泛函分析. 北京: 高等教育出版社, 1994.

[3] 复旦大学. 数学物理方程. 北京: 高等教育出版社, 1982.

[4] 冯果忱, 黄明游主编, 刘停战等编. 数值分析 (上册). 北京: 高等教育出版社, 2007.

[5] 冯果忱, 黄明游主编, 周蕴时等编. 数值分析 (下册). 北京: 高等教育出版社, 2007.

[6] 徐绪海, 朱方生. 刚性微分方程的数值方法. 武汉: 武汉大学出版社, 1997.

[7] 黄明游, 刘播, 徐涛. 数值计算方法. 北京: 科学出版社, 2005.

[8] 黄明游. 发展方程的有限元方法. 上海: 上海科学技术出版社, 1988.

[9] 李荣华. 偏微分方程数值解法. 北京: 高等教育出版社, 2005.

[10] 李荣华. 边值问题的Galerkin有限元法. 北京: 科学出版社, 2005.

[11] 胡建伟, 汤怀民. 微分方程的数值方法. 北京: 科学出版社, 1999.

[12] 郭本瑜. 偏微分方程的差分法. 北京: 科学出版社, 1988.

[13] 向新民. 谱方法的数值分析. 北京: 科学出版社, 2000.

[14] 马驷良, 李荣华. 关于矩阵族一致有界的代数准则 (综合报告). 吉林大学自然科学学报, 1986 (1): 21-36.

[15] 矢嶋信男, 野木达夫. 发展方程的数值分析. 王宝兴, 等译. 北京: 人民教育出版社, 1982.

[16] Samarski A A., Andreev V. B.. 椭圆型方程数值方法. 武汉大学计算数学教研室译. 北京: 科学出版社, 1984.

[17] Ciarlet D G. 有限元素法的数值分析. 蒋尔雄, 等译. 上海: 上海科学技术出版社, 1978.

[18] Hachbusch. 多重网格法. 林群, 等译. 北京: 科学出版社, 1998.

[19] Bramble James H. Multigrid methods, Pitman research notes in mathematics series, 294. Longman scientific & technical, 1993.

[20] Courant R, Hilbert D. Methods of mathematical physics, vol Ⅱ, J. wiley & Sons, 1962.

[21] Ciarlet D G. The finite element method for elliptic Problems. North Holland, Amsterdam - New York Oxford, 1978.

[22] Gear C W. Numerical initial value problem in ordinary differential equation. Prentice - Hall, 1971.

[23] Henrici P. Discrete variable methods in ordinary differential equations. J. Wiley & Sons, 1962.

[24] Lembert J D. Computional methods in ordinary differential equations. J. Wiley & Sons, 1973.

[25] Mihlin S. G. Variational method in mathematical physics, Moscow, 1957 (俄文).

[26] Marchuk G I. Methods of numerical mathematics. 2 Edition, Springer - Verlag, 1981.

[27] Quarteroni A., Valli A. Numerical approximation of partial differential equations, Springer Verlag, New York, Inc, 1997.

[28] Richtmyer R D, Morton K W. Difference methods for initial - value problems, 2nd edition. John wiley & sons, 1967.

[29] Strang G, Fix G J. An analysis of the finite element method. Prentice - Hall, 1973.

[30] Saua'yev V. K. Difference methods for solving equation of parabolic type, Moscow, 1960.

[31] Thomas J. W. Numerical partial differential equations, Springer - Verlag, New York, Inc, 1999.

[32] Varga R S. Matrix iterative analysis, Prentice - Hall, 1962.



[General Information]
书名=微分方程数值解法
作者=
页数=
出版社=
出版日期=
SS号=12186625
DX号=000006694728
url=http://www.jingjilei.cn/DrsPath.do?kid=686968686F6D6D696C3136343337323537&username=qhdx&spagenum=1&pages=50&fid=1156123 6&a=a6f0919f8f7cb70ee007149df5f61945&btime=2012-07-22&etime=2012-08-11&template=bookdsr1&firstdrs=http%3A%2F%2Fbook.duxiu.com%2FbookDetail.jsp%3FdxNumber%3D000006694728%26d%3D6EBD7F482869C881F08518379D0FB84E